Face Recognition Ability in Preterm or Low Gestational Weight Adults and Children

Maddie Atkinson

A thesis submitted in partial fulfilment of the requirements for the degree of Masters by Research (MRes), awarded by Bournemouth University

July 2018

Copyright statement

Face recognition ability in preterm or low gestational weight adults and children

Maddie Atkinson

Face recognition impairments can present throughout life as a result of acquired or developmental influences. While existing evidence implicates genetics and early visual deprivation, little is known about how other early influences may impact the development of the face recognition system. Very recent evidence suggests that premature birth and low birth weight influence face recognition ability in later childhood (Perez-Roche et al., 2017), however, the trajectory of these impairments is unclear. The present research aimed to address how the early influence of prematurity and/or low birthweight affects the trajectory and plasticity of the face processing system from childhood to adulthood. In Experiment 1, adults (n = 94) completed four dominant tests of face and object recognition ability to assess their relevant perceptual and mnemonic skills, completed measures of social functioning, and provided information on their birth weight and gestation. In Experiment 2, we monitored the eye-movements of a subset of these participants (*n* = 32) while they viewed a set of static images of people engaged in naturalistic social scenes, to detect any atypicalities in the face-processing strategy itself. Correlational analyses revealed that percentile (a combination measure of birth weight and gestation) was related to, and predicted, only face perception scores. The present results are unlikely to be accounted for by general perceptual processing mechanisms and co-occurring socio-developmental disorders. We also did not observe convincing evidence for reduced optimum processing with reduced face perception skill, suggesting that abnormalities in the face processing strategy

itself do not necessarily underpin atypical face perception skills. Most importantly, the present research suggests that, and at least in some cases, face-selective perceptual deficits remain consistent and persist from childhood (Perez-Roche et al., 2017) into adulthood. What remains unclear, as well as the theoretical and practical applications of this finding are discussed.

Table of contents

Table of figures

Table of tables

Acknowledgements

Firstly, I would like to express my gratitude to the Faculty of Science and Technology at Bournemouth University for offering me a scholarship to undertake this important and timely research, which built upon and extended a recent study by Perez-Roche et al. (2017); without funding, this project would not have been possible.

My first acknowledgement goes to those who make up my supervisory team. I would like to thank their collective knowledge for making this project possible: Dr Sarah Bate for formulating and creating the present research project; Dr Peter Hills for providing me with much-appreciated statistical guidance; and Dr Natalie Mestry for her continued support and constructive critiques that have allowed me to develop as an analytical and independent researcher throughout the year. Under their guidance I successfully overcame many difficulties and have learnt a lot that I will apply to my future projects as I progress on to the next stage of my research career.

In addition, I would also like to thank Dr Matthew Green, Ebony Murray, and Amanda Adams for their never-ending patience, support, and advice as I got to grips with the eye-tracking software. I also owe a thank you to Professor Stephen Powers and Dr Sarah Usher from Rothamsted Research Centre for their thoughtful feedback on various aspects of my research.

Thanks also goes out to all the individuals who took part in this research, and my friends and relatives who provided me with moral support and encouragement through difficult times.

1. Introduction

1.1. Individual differences in face recognition ability

The ability to recognise faces varies substantially between individuals and across different populations. In studies of the typical population, it is generally accepted that the (self-) reported and observed variation in face recognition ability can be (at least partly) attributable to genetics (e.g. Wilmer et al., 2010), personality traits such as extraversion (Li et al., 2010) and neuroticism (Perlman et al., 2009), and levels of socio-emotional functioning (Bate, Parris, Haslam, & Kay, 2010). Face-specific processing strategies, such as the ability to process faces as a whole, rather than their constituent parts (i.e. holistic processing) also predict face recognition ability (Wang et al., 2012; Richler, Cheung, & Gauthier, 2011). Less substantial evidence also suggests that variation in habitual sleep duration (Mograss, Guillem, & Stickgold, 2010) and adaptive coding of face identity (Rhodes et al., 2014) correlate with face recognition ability.

Although the interactive effects of influences on face recognition ability are unknown, most individuals report and exhibit successful face recognition skills. Some individuals, however, do experience significant impairments in face recognition ability, and are thought to have developmental prosopagnosia (DP), in which the ability to recognize faces simply fails to develop (Bate, 2013, p.59). Less severe reports of face recognition deficits suggest that other individuals may present only sub-clinical impairments. It is possible that these difficulties are underreported as individuals may often appear to exhibit adequate face recognition skills but use effective compensatory strategies to

recognize others (e.g. using social group status has been reported to be a successful compensatory processing strategy in older adults and has been related to underlying scanning patterns; Firestone, Turk-Browne, & Ryan, 2007). Further, research has identified a group of individuals who appear to have extraordinary face recognition abilities; these "super-recognizers" significantly out-perform controls on multiple tests of face recognition ability and are thought to be as good at face recognition as DPs are bad (Russell, Duchaine, & Nakayama, 2009). Taken together, these findings suggest that face recognition ability can be measured on a continuum; critically, if this is true, there may be observable factors that predict an individual's face recognition ability (Bate et al., 2010). While existing evidence indicates that face recognition ability is influenced by genetics (Wilmer et al., 2010) and early periods of atypical visual experience (e.g. Geldart et al., 2002), little is known about what other early influences may impact individual differences in face recognition ability.

Existing evidence indicates that adverse perinatal experiences, including prematurity and low birth weight, usually lead to atypical outcomes throughout childhood that often persist into adulthood. Cross-sectional (e.g. Stein, Siegel, & Bauman, 2006), longitudinal (Elgen, Sommerfelt, & Markestad, 2002), and meta-analytic (Bhutta et al., 2002; Aarnoudse- Moens et al., 2009) data show consistent and significantly worse behavioural and cognitive outcomes for low birth weight children when compared to their typical counterparts (for a systematic review, see Linsell, 2017). For example, children born very preterm perform significantly worse on tests of mathematics and reading ability than

children born full term when assessed at eight years of age (Anderson & Doyle, 2003). A higher prevalence of attentional deficits and internalizing behaviour problems, such as anxiety, have also been reported in this population (Indredavik et al., 2004). Similar observations have been documented over time (i.e. in adult samples; e.g. Hack et al., 2002) and in cohorts across countries (Saigal et al., 2003). It has also been proposed that the development of higher order visual abilities can be damaged by adverse perinatal factors such as prematurity or low birth weight (Pueyo et al., 2012). Very recent evidence raises the possibility that these factors may also influence face recognition ability in later childhood (Perez-Roche et al., 2017). In this study, children (aged 5-15 years) with adverse perinatal backgrounds (i.e. small for gestational age) were recruited through ophthalmology clinics and compared with matched controls. Small/appropriate for gestational age is a combination measure, termed percentile, that is customarily reported by other researchers (e.g. Figueras et al., 2008) which encompasses the influence of both birth weight and gestation. All children completed one face recognition memory test; children born small for gestational age scored significantly worse than children born appropriate for gestational age and were more likely to exhibit immediate and delayed face recognition memory deficits. Critically, the differences between small for gestational age and appropriate for gestational age children remained constant throughout childhood because the scores between the two groups did not disappear with age. However, the precise trajectory of these impairments remains unclear and it is unknown if they persist into adulthood or are merely delayed.

1.2. Face recognition ability over time

The results reported by Perez-Roche et al. (2017) imply multiple plausible trajectories of face recognition ability (see *Figure 1* for an illustration of these trajectories). The trajectory most consistent with the preceding literature is that early deficits impair the development and specialization of the face-recognition system through to adulthood. This trajectory is also supported by the Perez-Roche et al. (2017) data, which suggests a persistent alteration of face recognition ability throughout childhood might remain into adulthood. Studies (with similar designs) in adolescents born preterm have found both a similar lack of improvement in executive functions (Luu, Ment, et al., 2011) into adulthood and an alternative "catch-up" function by adolescence in cognitive and receptive vocabulary (Luu, Vohr, et al., 2011). It is possible that face recognition will follow the same trajectories as other general executive and cognitive functions. However, this is difficult to predict as face recognition is thought be a highly heritable, *specific* cognitive ability (Wilmer et al., 2010; see section *1.4.*), and studies of individuals with DP are consistent with a dissociation account for face and word processing, so it not necessarily the case that face recognition ability will follow the trajectory of receptive vocabulary (Burns et al., 2017). At present, it is unknown which trajectory face recognition ability will follow.

It is also plausible that early deficits merely delay the development of normal face recognition ability. Critically, if individuals "outgrow" all, or some, of their face recognition deficits, one question that arises is at what point in development does this occur; the Perez- Roche et al. (2017) study implies that if

present, this function would occur at some point after fifteen years of age. One study investigating the developmental trajectory of DP suggests the possibility that face perception deficits can improve or recover prior to, or during, adulthood (Dalrymple, Garrido, & Duchaine, 2014). However, there have been no published cases of individuals who experienced developmental memory and perception face recognition difficulties as a child, and not as an adult. It remains unclear whether face memory deficits can also improve over time. Alternatively, any developmental delays may reflect a sub- clinical population who are not as severely affected or who have developed effective compensatory strategies that are not successfully teased apart by current tests of adult face recognition ability. In other words, by the time these individuals reach adulthood, their performance is comparable to individuals in the typical population, but not optimal (perhaps due to divergent processing mechanisms).

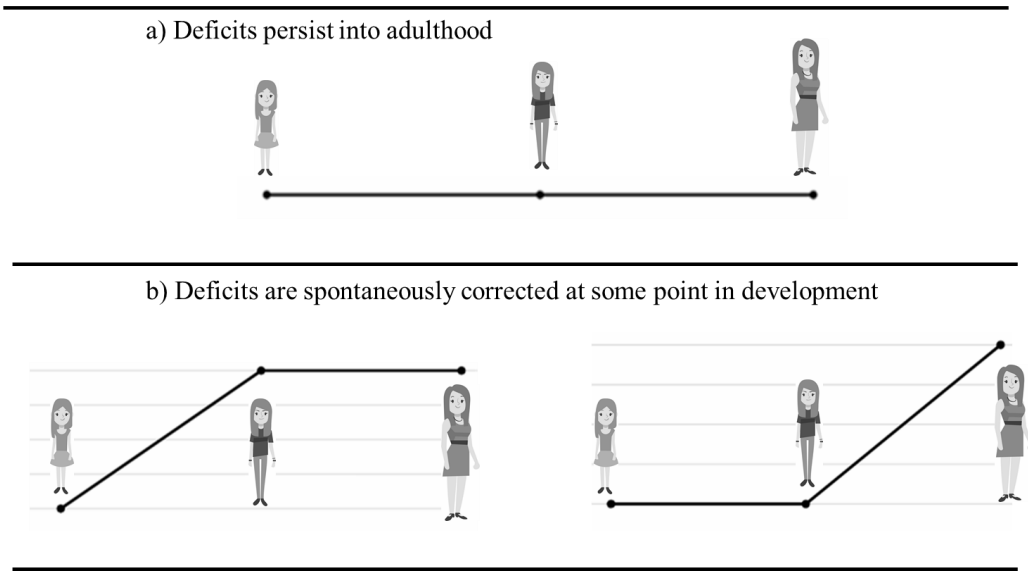

Figure 1. Possible developmental trajectories of face recognition ability for illustrative purposes. (a) Early deficits in face recognition ability are consistent and persist into adulthood; (b) early deficits can improve across development

or be spontaneously corrected at some point during development.

1.3. Potential mechanisms

Although recent (and preliminary) evidence has raised the possibility that prematurity and/or low birthweight (i.e. small for gestational age; Perez-Roche et al., 2017) can have a negative impact on face recognition ability in later childhood, the mechanisms that might be driving these atypicalities are not well understood. It has been proposed that the developmental changes in various cognitive abilities result from differences in structural and functional brain maturational processes between individuals born preterm or at-term (Mento & Bisiacchi, 2011; Johnson, 2010). Given that there are dramatic growth changes in foetal white matter volume, grey matter volume, and deep subcortical structures during the second and third trimesters (18 – 40 weeks), it is possible preterm birth may simply disrupt this process of structural organisation (Andescavage et al., 2016). This possibility is corroborated by structural MRI studies that provide evidence for abnormalities in specific brain regions and processes (Fenoglio, Georgieff, & Elison, 2017). For example, grey matter volume in the occipital face area, which has been implicated in accurate face perception (e.g. Pitcher, Walsh, & Duchaine, 2011), is significantly reduced for preterm infants in comparison to their term counterparts (Thompson et al., 2006). Further, the degree of prematurity has been associated with volume reduction in other brain regions, such as the orbitofrontal lobe (Ball et al., 2011), suggesting that there could be linear differences in face recognition abilities between moderately, very, and

extremely preterm individuals. Research has also shown that even at corrected age, preterm and term-born infants present differential cortical maturation for face recognition processes (Frie, Padilla, Ådén, Lagercrantz, & Bartocci, 2016).

In addition to altered cerebral maturation as a result of different prenatal and postnatal experience (i.e. gestational age), premature birth is associated with an increased risk of perinatal injuries, such as congenital cataracts. Studies investigating visual function in preterm and low birthweight individuals have found permanent deficits in visual acuity, contrast sensitivity, and colour detection (e.g. O'Connor et al., 2004). Such deficits may, in turn, result in a heightened risk of undergoing developmentally altered neurocognitive trajectories (Allen, 2008). To illustrate, individuals who have been deprived of early visual input as a result of bilateral congenital cataracts exhibit impaired performance on tasks of face recognition, but not on tasks of expression recognition, gaze direction, and lip speech (Geldart et al., 2002), suggesting that early visual deprivation can alter some aspects of face processing ability. Likewise, it is possible that if individuals can encode early visual information, they may simply be less efficient at processing such information. Preterm infants have been shown to be slower at encoding information than their term-born counterparts during the first year of life (Rose, Feldman, & Jankowski, 2001), suggesting that information processing may mediate the effects of prematurity and/or low birthweight on later cognitive abilities. Information processing abilities may be the foundation of later cognition; specifically, the "cascade" model suggests that elementary

infant abilities (such as information processing) influence higher-order abilities which, in turn, indirectly influences later cognitive abilities (Rose, Feldman, Jankowski, & Van Rossem, 2005). Thus, if early perceptual experiences are not optimal, functionally related pathways may not be collecting information in an efficient way (i.e. there may be poor neural connectivity between brain regions).

The myelination of white matter tracts and neural connectivity between brain regions are two factors that contribute to general brain development. As these processes begin in the third trimester (i.e. starting at 28 weeks), preterm birth might also disrupt this process; similarly, the pruning process may be less successful in preterm individuals (Innocenti & Price, 2005). White matter tracts have been strongly implicated with preterm brain development, as prematurity is associated with increased risk of specific brain insults (Fenoglio et al., 2017). Although MRI studies show that the anterior brains regions (which link mnemonic and affective information to faces) and posterior brain regions (which are responsible for basic perceptual function) are interconnected by long-range white matter tracts, it is less clear whether variation in the connectivity of these pathways explains individual differences in face recognition (Unger, Alm, Collins, O'Leary, & Olson, 2016). Specifically, the inferior longitudinal fasciculus (IFL) has been shown to predict variability in performance on valid and standardised tests of face memory, including the Cambridge Face Memory Test (Duchaine & Nakayame, 2006b). Similarly, research has demonstrated a robust negative association between the reduction of IFL integrity and the age-related changes

observed in face perception (Thomas et al., 2008). Consistent with this hypothesis, certain parts of the ILF have been associated with recognition ability for different classes of objects (Tayor et al., 2013). In this study, face recognition was highly associated with only the anterior part of the ILF in the right hemisphere, whereas place recognition was associated with the middle and posterior part of the IFL bilaterally. Evidence from outside the typical literature also supports this hypothesis. Adults with DP express an atypical structure-behaviour relationship near face-selective regions of the brain, suggesting white matter atypicalities are specific and localised to such regions, not to entire fasciculi, which in turn may have selective behavioural manifestations (Gomez et al., 2016). Together, this research demonstrates that face recognition abilities are related to white matter tracts in typical and atypical populations. Relatedly, epigenetic regulation and variations (i.e. methylation) might be a potential mechanism through which adverse perinatal experiences indirectly contribute to brain development in specific areas, which in turn could be associated with higher-order cognitive abilities, such as face recognition (Fumagalli et al., 2018).

Of course, understanding how and why face recognition deficits may occur in children born prematurely and/or underweight depends in part upon theoretical issues, such as whether face recognition is a specific ability resulting from a unitary or modular system, comprised of separable processes, such as memory and perception (see sections *1.4* and *1.5* for an exploration of these theoretical issues). It has been proposed that the specialisation of face recognition occurs progressively over time, supporting a gradual process of

modularisation, not pre-specified modules (Karmiloff-Smith, 1994). Here, we have proposed three potential mechanisms that might be driving the atypicalities in preterm individuals, as follows: structural organisation; increased risk of perinatal visual injuries; and functional interconnectivity between brain regions. To conclude, the maturational changes that occur during the second and third trimester may be interrupted or altered by preterm birth. As such, it is likely that, as a population, children born prematurely are at heterogenous risk of brain injury or insult (Briscoe, Gathercole, & Marlow, 2001). The preceding proposed mechanisms are by no means exhaustive and, collectively, suggest that further consideration of the interaction(s) between brain function, brain connectivity, and visual processing are necessary in order to understand their role in producing cognitive deficits in childhood through to adulthood.

To reiterate, the key claim of Perez-Roche et al.'s (2017) paper is that children born small for gestational age exhibit significant and persistent face recognition deficits in comparison to their appropriate for gestational age counterparts, and that these deficits may persist into adulthood. Although timely, these preliminary findings do not present convincing evidence of adverse perinatal experiences on face recognition ability because the authors fail to address some key theoretical issues, including the developmental relationship between faces and other classes of visual stimuli (i.e. non-face objects) and the interaction of visual memory and perception. Accordingly, the face-specificity of face recognition impairments and the dissociation between face memory and face perception will now be explored.

1.4. Faces versus objects

A question that speaks to a theoretically important debate is whether face and object processing follow the same developmental trajectories and to what extent face recognition impairments are face-selective. Multiple lines of evidence provide support for the notion that face recognition is 'special' and involves domain-selective cognitive and neural processes. For example, behavioural studies have shown that the cost of inverting faces is greater than that for other classes of mono-oriented objects (the "face inversion effect|"; e.g. Yin, 1969). Face recognition is also evidenced to be more accurate when faces are viewed as a whole, rather than separate parts (the "part-whole" effect); an effect often reduced (or absent) for non-face objects (e.g. Tanaka & Farah, 1993). These behavioural markers of 'holistic processing have also been shown to associate with brain activity (Yovel & Kanwisher, 2004 and Schiltz & Rossion, 2006, respectively). Neuroimaging studies show increased activation in the Fusiform Face Area (FFA) in response to faces, relative to a variety of non-face stimuli in typical populations (e.g. Kanwisher et al., 1997). In contrast, a range of non-face objects (such as dogs, birds, and cars) have shown comparable inversion effects (e.g. Gauthier, Skudlarski, Gore, & Anderson, 2000) and neural activation signatures (e.g. Xu, Liu, & Kanwisher, 2005) to faces, in individuals with sufficient expertise, suggesting that faces are not 'special', but that the selective processing is used only for objects of expertise. supporting a domain-general account. Critically, this account predicts that face and object recognition deficits always co- occur and that non-face object

recognition deficits are constant across categories. Such predictions are incompatible with the presence of face-selective cases of acquired (e.g. Busigny et al., 2010) and developmental (e.g. Duchaine & Nakayama, 2005) prosopagnosia, and an individual with DP who had expert-level within-class discrimination for one non-face object category (i.e. horses; Weiss, Mardo & Avidan, 2016).

Additional support for the face-specificity hypothesis comes from one case report of an individual demonstrating severe developmental object agnosia with intact face recognition abilities (Germine, Cashdollar, Duzel & Duchaine, 2011a), suggesting the possibility of a double dissociation. Importantly, any co-occurrence of object agnosia with face recognition deficits does not undermine the theory that faces are processed in a different way to other visual classes of objects. Co-occurrence may reflect independent systems that rely on shared mechanisms that developed in an atypical way (Garrido, Duchaine, & DeGutis., 2018), or common genetic and environmental factors that cause individuals to become vulnerable to multiple neurodevelopmental conditions (Gray & Cook, 2018).

One possibility is that selective deficits for faces (Ramus, 2004) result from focal, rather than extensive, developmental atypicalities affecting either face-processing regions or more generalized cortical areas, respectively. Face and object recognition impairments have been shown to associate in severity (e.g. Zhao et al., 2016), so it follows that individuals with both face *and* object deficits may have more widespread neural atypicalities. However, the linearity of this relationship, and what underlies more (or less) distributed atypical

development patterns, remains unclear. Relatedly, if face and object processing mechanisms are separable, a key theoretical question is at what stage of development does this separation occur. There is also a long-standing debate about the developmental trajectory of face recognition in typically developing children. While early studies suggest face-processing is qualitatively different in childhood and adulthood; that is, until ten years of age children process faces in parts rather than holistically (the "encoding switch hypothesis; Diamond & Carey, 1977), more recent studies suggest face processing *is* qualitatively adult-like by five years of age (Crookes & McKone, 2009). Partial resolution for these conflicting views suggests that memory (but not perception; this dissociation is addressed in section *1.4.*) processes develop at different rates for faces and objects, with steeper developmental slopes for faces (Weigelt et al., 2014). As such, we do not necessarily expect the trajectories to be the same for faces and objects. Perez-Roche et al.'s (2017) assessment battery did not include non-face object recognition tests, presumably because there are no object tasks parallel to the Facial Memory Subset of TOMAL (Test of Memory and Learning; Reynolds & Voress, 2007), so it is unclear whether the recognition impairments reported in this study are face-selective, or generalise to other object categories (i.e. non-face objects), across development. It is possible that low birth weight leads to more extensive and generalized developmental atypicalities and/or vulnerability to develop multiple neurodevelopmental conditions. If this is the case, there may also be a greater prevalence of object recognition deficits in individuals who have experienced a low birth weight than those who have experienced a typical birth weight.

In sum, the precise relationship between typical and atypical face and object recognition remains unclear, however many researchers subscribe to the belief that there are distinct cortical regions dedicated to face-processing (Duchaine & Yovel, 2015; Garrido et al., 2018; Gray & Cook, 2018; Rosenthal & Avidan, 2018). Another possibility is that perception and memory related processes impact face and non-face object categories differently (Towler & Tree, 2018). Thus, to be informative regarding the face-selectivity and developmental trajectory of face recognition impairments, studies should include both face memory and face processing tasks, along with object parallels of these tasks (Starrfelt & Robotham, 2018).

1.5. Memory versus perception

Although the term face *recognition* has often been used within the literature to refer to either face perception or face memory, models of face-processing propose that these processes fall in to discrete cognitive stages. Face perception refers to the ability to discriminate faces, without a memory component. On the other hand, face memory refers to the ability to commit individual faces to memory and recall them. Although this latter process does rely heavily on face perception, it also requires additional processes, such as conscious awareness that the face has been encountered before (Weigelt et al., 2014). Specifically, Bruce and Young's (1986) seminal model outlines a separation between the structural encoding of a face (analogous to face perception) and face recognition units, which encode faces in long-term memory (face memory). This sequential model proposes that face recognition

impairments can result from failures at one or more stages of processing. According to this model, face memory requires face perception but face perception does not require face memory. Although it is face memory, and not face perception, that more closely mimics facial identity recognition in everyday life (Duchaine & Nakayama, 2005), both processes are necessary for successful face recognition. The literature supports a developmental dissociation between face perception and face memory. In typical populations, face perception appears to mature in early childhood and at the same rate as perception for other classes of objects, whereas face memory follows a protracted period of development, until mid-adolescence, that diverges from other classes of objects (Weigelt et al., 2014).

Evidence from outside the typical literature also suggests that face perception and face memory may engage partly dissociable mechanisms, although we are not expecting small for gestational age individuals to be clinically impaired at face recognition. Case studies of DP detail individuals who are impaired at both face perception and face memory (e.g. Duchaine & Nakayama, 2006b) and individuals who are impaired at only face memory (e.g. Palermo et al., 2011). Studies that report both behavioural and implicit measures (such as reaction times; RTs) suggest the possibility that 'typical' performance (i.e. accuracy scores) reflects the successful application of compensatory strategies (Duchaine & Nakayama, 2004). For example, adult DPs exhibiting 'intact' face perception scores are reliably slower at standardized tasks than control participants (e.g. Behrmann et al., 2005).

A recent study has demonstrated a dissociation between face memory

and face perception in adults, but not children with DP (Dalrymple, Garrido, & Duchaine, 2014). Interestingly, the adults who scored in the typical range for accuracy had normal inversion effects and RTs, suggesting that they did not use atypical or divergent processing mechanisms. Although Weigelt et al. (2014) propose that the typical development of face perception is mature in children as young as five years of age and follows the same trajectory as other classes of objects, it may be that children with face recognition deficits show delayed development of face perception. Dalrymple et al.'s (2014) paper lacks some explanatory value because tasks were not matched across categories (i.e. face and non-face objects). However, these findings raise the intriguing possibility that face recognition impairments may be qualitatively different in childhood and adulthood and that the process responsible for the development of face perception is merely delayed in some, but not all, cases of DP in children. Likewise, adverse perinatal factors (and potentially their severity) may impact perceptual processes differently, given that impairments can result from atypical face perception skills, face memory skills, or both.

Given that the typical development of face-processing skills relies on both memory and perceptual mechanisms, it is especially important to measure each process, without confounding the two. Face-processing skills are often assessed using the Cambridge Face Memory Test (CFMT; Duchaine & Nakayame, 2006b) and the Cambridge Face Perception Test (Bowles et al., 2009) in tandem. Additional support that memory and perception rely, at least in part, on separable mechanisms comes from examining the correlation of participant performance (as indexed by the CFMT and the CFPT) between the

two. Although the performances on the two tasks are highly correlated across samples of the typical population (approx. r = .61; Bowles et al., 2009), the fact that the correlation is not perfect (i.e. r = 1) suggests that face memory and face perception are partially dissociable skills. As previously mentioned, some individuals with DP to show intact face perception but poor memory skills (e.g. Palermo et al., 2011). It is also possible that some DPs can achieve 'typical' scores on the CFMT by utilizing effective compensatory strategies (Duchaine & Nakayama, 2004). Interestingly, a meta-analysis of 90 studies that investigated face processing impairments in Autism Spectrum Disorder (ASD) showed that face-processing impairments were only revealed when the assessment task involved memory demands, not when the tasks were perceptual in nature (Weigelt et al., 2012). Taken together, converging evidence supports the notion that face memory and face perception processes are supported by separate cognitive and neural systems.

Overall, Perez-Roche et al.'s (2017) use of only the Facial Memory Subset (a test of face memory) was unable to address many of the key theoretical questions outlined in the preceding literature. Specifically, it remains unclear whether the impairments reported in Perez-Roche et al.'s (2017) study result from selective failures of face perception, and whether these impairments are permanent, or merely represent a developmental delay. Likewise, the use of only one test (irrespective of its suitability) cannot differentiate between individuals who may have performed within typical range (possibly by chance) but are (sub)clinically impaired at face recognition, and those who may have performed below the typical range but do not have face recognition impairments

(Bate & Tree, 2017). This highlights the importance of using a 'multi-test' approach, including not only behavioural tests, but also different measures of face and object processing. For example, eye tracking studies can be useful to assess how individuals visually explore faces and objects and can be used identify atypicalities in the face processing strategy itself (Eimer, 2018).

1.6. Alternative methodologies

One way of investigating the underpinnings of face recognition impairments is through the analysis of individual eye movement strategies. Despite similar scores across behavioural measures, adult DPs with impaired face memory can display typical face perception in terms of error rates, even though they are significantly slower at the perception tasks than matched controls (Behrmann et al., 2005), suggesting the use of atypical or divergent processing mechanisms. These comparable levels of accuracy suggest that DPs can often 'get by' through the use of compensatory strategies that are successful at least some of the time (Bate et al., 2015; Bate & Bennetts, 2014; Duchaine & Nakayama, 2004). Similarly, individuals with sub-clinical deficits are likely to use successful compensatory strategies, and often appear to exhibit adequate face recognition abilities (e.g. Firestone et al., 2007). The examination of individual eye-movement strategies provides an online measure of cognitive processing which can reveal processing strategy in a way that accuracy and reaction time data cannot (e.g. Liversedge & Findlay, 2000).

It is thought that successful face recognition is likely to depend on specific fixation patterns, and there is considerable evidence to suggest that the eye region (e.g. Taylor et al., 2001; Gosselin & Schyns, 2001), and the nose

region, (e.g. Hsiao & Cottrell, 2008) are critical for face identification. Studies that implicate the eye region have examined both the eyes together (as a single region; e.g. Peterson & Eckstein, 2012) and the eyes separately (the left eye and the right eye, as two distinct regions; e.g. Vinette, Gosselin, & Schyns, 2004). Both 'types' of analysis converge on the same conclusion that the eyes are (one of) the most important facial features in facial recognition. However, when the eyes are considered separately, the left eye has been shown to be used more effectively and more rapidly than the right (Vinette et al., 2004). Earlier research also supports a bias for the left half of a face (e.g. Burt & Perret, 1997). Further, it has been demonstrated that the left eye drives the N170 effect just as well as both the eyes in the right hemisphere; the converse is true in the right hemisphere (Smith, Gosselin, & Schyns, 2004). Thus, if the left eye is attended to before the right, it is possible that the right hemisphere of the brain processes faces more efficiently than the left. This concept is supported by the fMRI literature, that has shown that the fusiform face area (FFA) in the left hemisphere is activated to a greater degree than the FFA in the right hemisphere (e.g. Kanwisher, McDermott, & Chun, 1997). As such, there might be differences between the left and the right eyes, which may also relate to neural processes.

In some DP cases, individuals have been shown to spend less time examining the eye region, and more time examining the mouth region, in comparison to controls; whereas individuals with severe DP have been shown to spend significantly less time examining these internal facial regions than controls when free- viewing social scenes (Experiment 1; Bobak et al., 2017).

The former finding suggests a quantitative difference, whereas the latter finding supports the notion that, in at least in some cases, DP is characterized by qualitative differences between populations. Given that the present focus is on sub-clinical impairments, one could reasonably expect a quantitative relationship between eye movement strategies and birth weight for gestational age. Moreover, it is possible that atypicalities in face processing may be detectable from birth, or instead manifest at a particular stage of development.

To conclude, eye-movement recording techniques may provide a more appropriate and reliable way of detecting atypical face recognition skills in individuals with sub-clinical deficits; alleviate some of the methodological caveats associated with interpreting behavioural data from face and object recognition tests; and can identify atypicalities in the face processing strategy itself that may or may not be face-selective (Schwarzer et al., 2007).

1.7. The current research

This large-scale ongoing research project aims to address how the early influence of prematurity and/or low birth weight affects the trajectory and plasticity of the face processing system from childhood to adulthood, using explicit (Experiment 1) and implicit (Experiment

2) methodologies. Age-appropriate versions of all tasks were administered to children at various stages of development, however, due to difficulties with recruitment leading to an incomplete child dataset at present, only the adult data are reported and interpreted in this thesis. Determining whether adverse perinatal experiences influence later face recognition ability could elucidate the developmental trajectories of both typical and atypical face

processing skills; have important implications for healthcare providers, caregivers, and parents of premature and/or low birth weight infants; and, in the longer term, may identify a group of individuals who would benefit from early intervention that addresses the development of face processing skills. In Experiment 1, dominant tests of face and object recognition were administered to adults to determine whether there is a (dis)similar relationship between face recognition ability and birth weight for gestational age between children (i.e. Perez-Roche et al., 2017) and adults. Given that behavioural and implicit measures do not often correlate (e.g. Behrmann et al., 2005), Experiment 2 employed a social scenes eye- tracking paradigm to detect any abnormalities in the face processing strategy itself among a subset of these adult participants. Specifically, the relationships between dwell times across the inner versus outer facial features, and across the eyes, nose, and mouth, and birth weight for gestational age were investigated.

1.8. Hypotheses

The overall aim of this project is to investigate the developmental trajectory of face recognition impairments that result from premature birth and/or low gestational weight. We aim to extend the findings of Perez-Roche et al. (2017). Our hypotheses and predictions were as follows:

(1) To investigate whether the perinatal effects on face recognition ability persist into adulthood. If perinatal effects do persist, we would expect to find that birth weight for gestational age is related to, and predicts, scores on face recognition tests. However, if face recognition impairments

are merely delayed, we would expect to see no relationship between birth weight for gestational age and scores on face recognition tests.

(2) To investigate whether these face recognition impairments represent general perceptual deficits or are face-selective. Face-selective impairments would be indicated by a significant relationship between birth weight for gestational age and face processing tasks, and no concurrent relationship on the matched object tasks; general impairments would be indicated significant relationships between birthweight for gestational age and both face and object processing tasks.

(3) To investigate whether perinatal effects on face recognition ability are limited to perceptual processes (i.e. Perez-Roche et al., 2017). The inclusion of dominant tests in this project will provide stronger evidence as to whether impairments result from selective failures of face perception, face memory, or both.

(4) To investigate whether behavioural tests of face recognition ability and eye-movement measures dissociate (i.e. that is, we observe a relationship on one measure but not the other). In the absence of an impairment in face recognition test scores, individuals who have experienced a low birth weight for gestational age may use an atypical face processing strategy. As this study is investigating sub-clinical deficits, we might

only expect to see quantitative relationship between birth weight for gestational age and dwell times. Specifically, a low birth weight for gestational age might be related to less time spent looking at regions of the face thought to be critical in successful face recognition, such as the eyes or nose, rather than a focus on different facial regions (i.e. a qualitative difference).

2. Experiment 1

Experiment 1's behavioural approach was designed to address the shortcomings of previous research. Specifically, we chose dominant tests of face memory and face perception and included object tasks matched in format and procedure to these tests to investigate whether face recognition impairments are face-selective and whether they relate to face memory processes, face perception processes, or both. Further, in order to combat alternative hypotheses, we also administered two questionnaires measuring socio-emotional functioning; participants who score beyond the cut-off were excluded from analysis, as a methodological control. Relatedly, we investigated whether there is a relationship between birth weight for gestational age and socio-emotional functioning, although this was not central to our overall aims.

2.1. Method

2.1.1. Participants

114 adults were recruited through posters at Bournemouth University and social media platforms (e.g. Twitter). Adults who provided incomplete data

(*n* = 12), had problems with their vision (i.e. illusory palinopsia, *n* = 1; however, participants who reported weak eyes or short-sightedness were included, *n* = 2), or had a diagnosed or suspected (i.e. participants were excluded if they scored less than 30 points on the Empathy Quotient or above 77 points on the Systemizing Quotient; Cohen & Wheelwright (2004) and Wheelwright et al. (2006), respectively) developmental or neurological disorder (*n* = 7) were excluded from analysis. This resulted in a total of 94 adults (70 = females, 82 = right- handed), aged 18-48 years (*M* = 23.49, *SD*= 5.84), in the final analysis. Reported birth weight ranged from 1kg to 4.86kg (*M* = 3.25, *SD* = 0.70). Birth weight was classified based on criteria set by UNICEF and the World Health Organisation (UNICEF & WHO, 2004): 82 participants were classed as a typical birth weight (>2.5kg; *M* = 3.45, *SD* = 0.48), 10 participants were classed as a low birth weight (<2.5kg; *M* = 2.09, *SD* = 0.33), and 2 participants were classed as a very low birth weight (<1.5kg; *M* = 1.10, *SD* = 0.13). Gestation period ranged from 25 weeks to 47 weeks (*M* = 39.13, *SD* = 2.70). Percentiles (a combination measure of birth weight and gestation) ranged from the 0th to 99th percentile (*M* = 44.71, *SD* = 30.47); 17 participants were below the 10th percentile (*M* = 2.41, *SD* = 2.81) and experienced a low weight for gestational age. Participants provided written consent and were compensated with either course credit or a £10 Amazon voucher for their time. Ethical approval for this study was granted by Bournemouth University's Ethics Committee (ethics application 17501).

2.1.2. Stimuli and materials

Below are brief descriptions of the memory tasks, perception tasks, and

socio- functioning measures used within this study. Participants were instructed to complete all tasks on either a laptop or a computer with a stable internet connection; work through each task one by one in the order that they were presented to them by clicking on each link within the email individually; position themselves in a quiet room; and complete all tasks in one sitting. Example stimuli are displayed in Figure 2.

Face memory task. Participants completed the CFMT+ (Cambridge Face Memory Test – Long Version; Russell et al., 2009), an extended version of the original CFMT (Duchaine & Nakayama, 2006b) which includes an additional 30 "hard trials". The stimuli in the CFMT+ are faces of Caucasian males. All faces display a neutral expression, images are greyscale and cropped to exclude external features (i.e. hair and ears) and paraphernalia (i.e. facial hair). The task consists of four phases: a learning phase; a test phase with novel images; a test phase with novel images degraded by visual noise; and a test phase with novel images that vary in pose and emotional expression and are heavily degraded using visual noise. In the learning phase, participants are shown one face from three different viewpoints for three seconds each. Participants are then presented with three faces (i.e. one target and two distractors) simultaneously and asked to pick out which identity they just saw, by pressing 1, 2, or 3 on the keyboard. After the learning phase, participants are presented with a review image of all the learned faces for 20 seconds. Participants then completed the first test phase, in which faces were presented in novel lighting conditions and/or viewpoints.

Following this test phase, participants were presented with a second

review image. Participants completed the second test phase, in which faces were presented from novel viewpoints and were degraded by visual noise. This was again followed by a revision period, and the final test phase of "hard trials". In this phase, the distractor images also recur more frequently than in the previous stages to minimise the difference in familiarity between the learned and distractor faces. The test included a total of 102 trials. Total scores (i.e. items correct out of 102) were calculated by summing the number of correct items from the four stages: in the learning phase, items correct out of 18; in the test phases, items correct out of 30, 24, and 30, respectively. A higher score indicates better performance.

Object memory task. Participants completed the CCMT (Cambridge Car Memory Test; Dennett et al., 2012). This task is matched in format to the original CFMT (Duchaine & Nakayama, 2006b), but the task uses images of cars rather than faces. All cars are the same colour and are edited to remove any obvious brand or design elements. The task consists of the same three experimental stages: a learning stage (18 trials); a test stage with novel images (30 trials); and a test stage with novel images degraded by the presence of visual noise (24 trials). Total scores (i.e. items correct out of 72) were calculated by summing the correct items from the three stages. A higher score indicates better performance.

Face perception task. Participants completed the CFPT (Cambridge Face Perception Test; Bowles et al., 2009). The stimuli in the CFPT are faces of Caucasian males. Again, faces have a neutral expression, and images are greyscale and cropped to remove external features and paraphernalia. On each

trial, participants have one minute to sort six faces from most-to-least similar to a target face. This task consists of 16 trials in total: eight with the faces upright and eight in an inverted format. For each trial, the perfect score is 0. Error scores for each item are computed by summing the deviations from the correct arrangement for each face (i.e. if a face is two positions from its correct arrangement, that is 2 errors). Error scores for the upright and inverted trials are calculated separately, and the maximum number of errors for the eight trials is 144. Proportion correct is calculated by subtracting a participant's error score from the maximum number of errors and dividing this difference by the maximum number of errors. A higher proportion correct indicates better performance.

Object perception task. Participants completed a car perception test (currently unpublished, Bournemouth University), matched exactly in format and procedure to the CFPT (Bowles et al., 2009), except that the stimuli are cars rather than faces. All cars were greyscale and edited to remove any obvious brand or design elements. Likewise, performance is measured as the total number of errors and scoring procedures mirror the CFPT (above). A higher proportion correct indicates better performance.

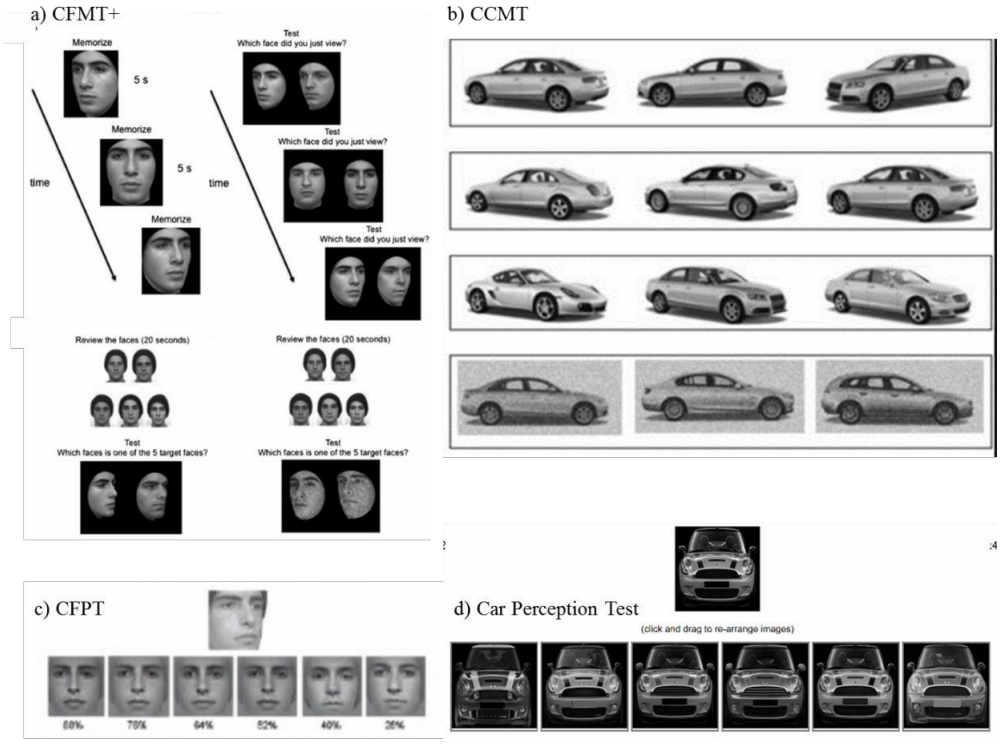

Figure 2. Example stimuli for a) the Cambridge Face Memory Test (CFMT+) taken from Duchaine and Nakayama (2006b); b) the Cambridge Car Memory Test (CCMT), taken from Dennett et al. (2012); c) the Cambridge Face Perception Test (CFPT), taken from Bowles et al. (2009); and d) the car perception test (currently unpublished, Bournemouth University).

Socio-emotional functioning measures. Participants completed two self-report questionnaires; the Empathy Quotient (EQ; Cohen & Wheelwright, 2004) and the Systemizing Quotient (SQ; Wheelwright et al., 2006). The EQ is designed to measure how quickly one might pick up on others' feelings and/or how strongly one is affected by these feelings. The questionnaire has 40 items and participants respond on a 4-point Likert scale from strongly agree to strongly disagree. An example item is "I can pick up easily if someone says one thing but means another". Each item can be scored with a maximum of 2 points; the maximum score on the EQ is 80 and the minimum is 0. The EQ has a test-

retest reliability of .97 and has been shown to relate to other measures of socio-emotional functioning (Cohen & Wheelwright, 2004). For the EQ, lower scores indicated greater autistic traits. The SQ is designed to assess an individual's drive to analyse, explore, and extract the underlying rules that govern a system. The EQ and SQ often dissociate and one individual may score typically on one measure, and not the other. The SQ is comprised of 75 items and participants respond on a 4-point Likert scale from strongly agree to strongly disagree. An example item is "when I look at a building, I am curious about the precise way it was constructed". Again, each item can be scored with a maximum of 2 points; the maximum score on the SQ is 150 and the minimum is 0. The SQ has also been shown to relate to other measures of socio-emotional function (Wheelwright et al., 2006). For the SQ, higher scores indicated greater autistic traits.

2.1.3. Design

There were three (continuous) independent variables: birth weight, gestation period, and percentile. Percentile was calculated for each participant prior to analysis using an online calculator (see http://www.paediatrics.co.uk/nicu/growth-charts). We chose to also calculate the percentile measure for three reasons. First, this measure is customarily reported by other infant researchers across the literature (e.g. Perez-Roche et al., 2017; Chen, Claessens, & Msall, 2014), Second, percentiles are useful for comparing values between infants (and studies) as it is a measure of relative standing in the population. Specifically, as percentile represents a combination of birthweight, gestation, and sex information, each score can be compared to

those of other children the same age. Third, because the percentile measure combines information from multiple sources, it takes account of each measure in relation to the other measures. Percentile allows individuals to be 'ranked'. For example, two individuals may share the same birthweight but differ in terms of their gestational age. Likewise, even though two individuals may share the same gestational period, one may be a very low birthweight and the other a typical birthweight. As such, if birthweight and gestation period are only measured independently, it does not give a full account of an individual's perinatal experiences. Participants provided scores on all four tests and two questionnaires, so there were six dependent variables in total: face memory scores (CFMT+), face perception scores (CFPT), car memory scores (CCMT), car perception scores (CCPT), Empathy Quotient scores, and Systemizing Quotient scores.

2.1.4. Procedure

Participants were contacted via email and completed all tasks online (consent forms and questionnaires were administered through www.qualtrics.com and the face and object recognition tests were administered through www.prosopagnosiaresearch.org). We considered the administration of all the tests online to be appropriate for two reasons. First, Web sampling made it more achievable to recruit a larger number of individuals who experienced a low birth weight for gestational age (i.e. a relatively rare population characteristic; 7.0% in England (Office for National Statistics, 2016)). And second, Web and lab-based samples have been shown to yield comparable data in terms of mean performance, performance variance, and internal reliability for

challenging cognitive tasks, such as the CFMT (Germine et al., 2012). Participants first signed consent forms to confirm their willingness to take part in the study. The tasks took approximately 1 hour to complete.

Memory tasks were always completed before perception (matching) tasks and face tasks were always completed before object tasks. Participants provided information on birth weight and gestation, then completed the face and object tasks, followed by the questionnaires (which were counterbalanced across participants).

2.2. Results

Reported birth weights were converted to kg, gestations were converted to weeks, and percentile was calculated by combining these two measures, using an online calculator. Scores on the four dependent variables of interest (i.e. CFMT+, CCMT, CFPT, CCPT) were converted to percentage correct prior to analysis. A Kolmogorov-Smirnov test was used to assess the sample distribution of all variables. Percentiles (a combination measure of birth weight and gestation) (D (94) = .071, p = .200), CFMT+ (D (94) = .060, p = .200) scores, CCMT (D (94) = .069, p = .200) scores, CCPT (D (94) = .085, p = .092) scores, and Empathy Quotient (D (94) = .055, p = .200) scores were normally distributed. Birth weight (D (94) = .127, p = .001), gestation (D (94) = .193, p <.001), CFPT (D (94) = .122, p =.001) scores, and Systemizing Quotient (D (94) = .115, p =.004) scores deviated significantly from normality. Transformations did not normalise the data, so non-parametric statistics are reported where appropriate.

Given the relative paucity of low birth weight individuals in the typical population (7.0% in England; Office for National Statistics, 2016), our group sizes were insufficient ($n_{lowbirthweight}$= 12, $n_{typicalbirthweight}$ = 82) for a between-groups analysis. Consequently, analyses were correlational in nature. Correlations were first conducted to assess the appropriateness of the combination percentile measure and to check that there were not issues with multicollinearity. Following this, separate multiple linear regressions were conducted to predict face and object recognition scores from each infancy-related IV.

2.2.1. Correlations

Correlational analyses were used to examine the relationship between participants' birth weight, gestation period, percentile scores (a combination measure of birth weight for gestational age) and their performance on four face and object recognition tasks and measures of socio-emotional functioning (see *Table 1*). Two-tailed correlations revealed a weak positive relationship between CFMT+[1] and birth weight (r_p = .207, p = .047) and CFPT and percentile (r_s = .275, p = .008). A weak positive relationship between CFPT and birthweight was approaching significance (r_s = .201, p = .054). Scatterplots were produced only for each of these relationships (see *Figure 3*). No other correlations were significant (all *p*s>.194). Higher birth weights and percentiles were associated with greater face memory and perception scores but were unrelated to object

[1] Note that the CFMT+ and the abridged CFMT are highly correlated with each other; rp = .960, p < .000. As such, the correlational results do not change *substantively* when using the abridged version of the CFMT, which is matched more closely to the CCMT (see Appendix A for CFMT results summary).

recognition (CCMT and CCPT) and socio-emotional functioning (empathy and systemizing quotients) measures.

Table 1. Correlation coefficients of birth weight, gestation, and percentile with face and object recognition scores and measures of socio-emotional functioning. p-values are reported in parentheses and significant correlations are highlighted in bold. Correlations indicating a trend towards significance are italicised.

	Hypothesised Predictors		
	Birth weight	**Gestation**	**Percentile**
CFMT+	**.207 (.047)**	.136 (.194)	.177 (.089)[a]
CCMT	.120 (.252)	.113 (.279)	.052 (.620)[a]
CFPT	*.201 (.054)*	.066 (.533)	**.275 (.008)**
CCPT	-.004 (.967)	-.094 (.369)	-.061 (.563)[a]
Empathy Quotient	.026 (.808)	.042 (.690)	.027 (.799)[a]
Systemizing Quotient	-.020 (.849)	-.036 (.732)	-.065 (.538)

Note. CFMT+ = Cambridge Face Memory Test – Long Version, CCMT = Cambridge Car Memory Test, CFPT = Cambridge Face Perception Test, CCPT = Cambridge Car Perception Test. Higher scores indicate better performance on all memory and perception tests, and greater autistic traits on both questionnaires.

[a] *Parametric coefficients reported.*

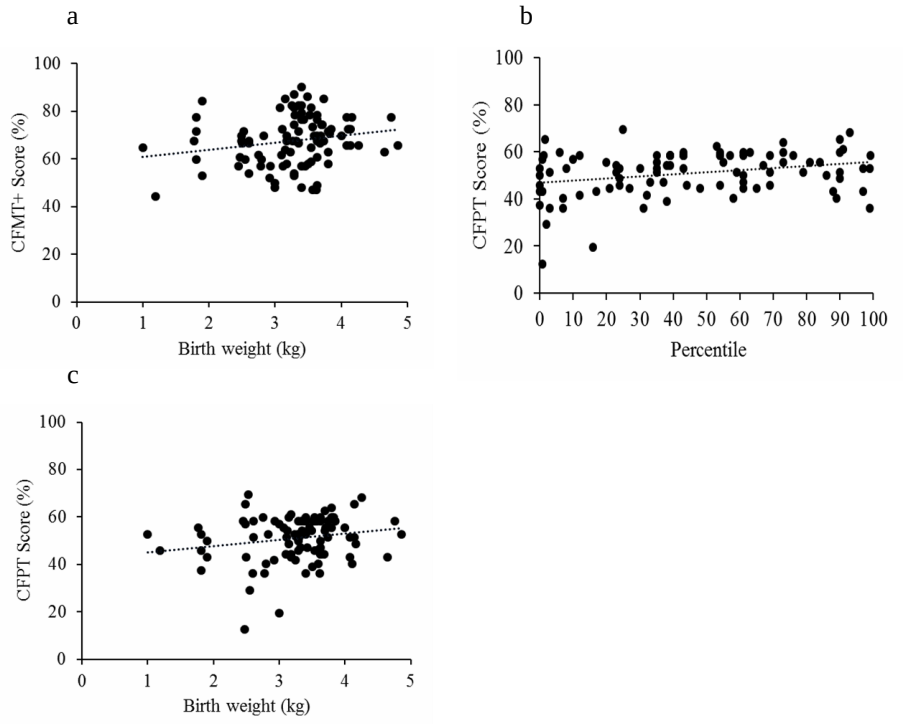

Figure 3. Scatterplots to show the relationship between a) birth weight and CFMT+ test scores; b) percentile and CFPT test scores; and c) birthweight and CFPT test scores. Coefficients signified only weak positive relationships, and the relationship illustrated by c) was only approaching significance.

As expected, all three infancy-related hypothesised predictors were positively related and correlated with each other (birthweight and gestation, r = .453; birthweight and percentile, r = .754; percentile and gestation, r = .103). Likewise, participants' scores across the various dependent measures were correlated with each other (see *Table 2* for correlation matrix). As these associations range from weak to moderately strong, multicollinearity is not a problem and results can simply be interpreted as they are (Goldberger, 1991; Kraha et al., 2012). Nevertheless, three separate multiple linear regressions for

each infancy-related IV were conducted.

Table 2. Correlation matrix of the dependent variables. Correlations ranged from weak to moderately strong.

	Dependent Measures					
	CFMT+	**CCMT**	**CFPT**	**CCPT**	**Empathy Quotient**	**Systemizing Quotient**
CFMT+	1	.311	.230	.030	.072	.058
CCMT	-	1	.282	.220	.135	.117
CFPT	-	-	1	.173	-.010	.029
CCPT	-	-	-	1	.126	.029
Empathy Quotient	-	-	-	-	1	.029
Systemizing Quotient	-	-	-	-	-	1

Note. CFMT+ = Cambridge Face Memory Test – Long Version, CCMT = Cambridge Car Memory Test, CFPT = Cambridge Face Perception Test, CCPT = Cambridge Car Perception Test. Higher scores indicate better performance on all memory and perception tests, and greater autistic traits on both questionnaires.

2.2.2. Multiple linear regressions

As percentile represents a combination measure of birthweight and gestation; and all three variables are correlated with each other (see *section 2.2.1*); and this measure is customarily reported by other researchers (e.g. Perez-Roche et al., 2017), all three variables were subjected to regression analyses.

Specifically, three separate multiple linear regressions were conducted to predict face and object recognition scores from each infancy-related IV: birthweight, gestation, and percentile. In each case, there is one continuous IV (i.e. birthweight, gestation, or percentile) and multiple DVs (CFMT+[2], CCMT, CFPT, CCPT, EQ, and SQ). As we are simply quantifying the relationships between variables, the infancy-related IVs were each separately entered into SPSS as a DV and the CFMT+, CCMT, CFPT, CCPT, EQ, and SQ scores were entered into SPSS as predictors. The following results are reported as entered into the model (i.e. face and object recognition scores and socio-emotional functioning measures predict infancy-related measures) but are interpreted as the reverse (i.e. infancy-related measures predict face and object recognition scores and socio-emotional functioning measures).

Birthweight

Table 3 provides a summary for the birthweight regression analysis. The model was non-significant; $F(6, 87) = 1.310$, $p = .261$, accounting for 2.0% of the variance (Adj. $R^2 = .083$). As presented in *Table 3*, all predictors were non-significant (all ps $\geq .101$). Birthweight did not predict face and object recognition scores or socio-emotional functioning measures.

[2] Note that the CFMT+ and the abridged CFMT are highly correlated with each other; $rp = .960$, $p < .000$. As such, the regression results for each IV do not change substantively when using the abridged version of the CFMT, which is matched more closely to the CCMT (also see *Appendix A* for CFMT results summary).

Table 3. Summary of the multiple linear regression analysis predicting birthweight from CFMT+, CCMT, CFPT, CCPT, EQ and SQ scores.

Variable	*t*	b	β	*F*	R^2	Adj. R^2	*p*
Overall Model				1.310	.083	.020	.261
CFMT+ Scores	1.655	.013	.188				.101
CCMT Scores	-.430	-.003	-.051				.668
CFPT Scores	1.658	.014	.187				.101
CCPT Scores	.500	.002	0.54				.618
EQ Scores	-1.221	-.008	-.132				.226
SQ Scores	.535	.004	.056				.594

Note. CFMT+ = Cambridge Face Memory Test – Long Version, CCMT = Cambridge Car Memory Test, CFPT = Cambridge Face Perception Test, CCPT = Cambridge Car Perception Test, EQ = Empathy Quotient, SQ = Systemizing Quotient.

Gestation Period

Table 4 provides a summary for the gestation period regression analysis. The model was non-significant; $F(6, 87) = 0.585$, $p = .741$, accounting for 2.7% of the variance (Adj. $R^2 = .039$). As presented in *Table 4*, all predictors were non-significant (all *p*s ≥ .210). Gestation did not predict face and object recognition scores or socio-emotional functioning measures.

Table 4. Summary of the multiple linear regression analysis predicting gestation period from CFMT+, CCMT, CFPT, CCPT, EQ and SQ scores.

Variable	*t*	b	β	*F*	R^2	Adj. R^2	*p*
Overall Model				0.585	.039	-.027	.741
CFMT+ Scores	1.262	.038	.146				.210
CCMT Scores	.367	.010	.044				.714
CFPT Scores	-.072	-.002	-.008				.943
CCPT Scores	.484	.009	0.53				.630
Empathy Quotient	-1.232	-.030	-.136				.221
Systemizing Quotient	-.219	-.003	-.023				.827

Note. CFMT+ = Cambridge Face Memory Test – Long Version, CCMT = Cambridge Car Memory Test, CFPT = Cambridge Face Perception Test, CCPT = Cambridge Car Perception Test, EQ = Empathy Quotient, SQ = Systemizing Quotient.

Percentile (combination measure of birthweight and gestation period)

Table 5 provides a summary for the percentile regression analysis. The model was non-significant; $F(6, 87) = 1.787$, $p = .111$, accounting for 4.8% of the variance (Adj. $R^2 = .110$). As presented in *Table 5*, CFPT scores was a significant predictor of percentile ($\beta = .246$, $t = 2.216$, $p = .029$); all other predictors were non-significant (all $ps \geq .245$). Percentile significantly predicted face perception scores but was unrelated to face memory scores, object recognition scores, and socio-emotional functioning measures.

Table 5. Summary of the multiple linear regression analysis predicting percentile from CFMT+, CCMT, CFPT, CCPT, EQ and SQ scores.

Variable	*t*	b	β	*F*	R^2	Adj. R^2	*p*
Overall Model				1.787	.110	.048	.111
CFMT+ Scores	1.171	.379	.131				.245
CCMT Scores	-.373	-.106	-.043				.710
CFPT Scores	2.216	.783	.246				.029*
CCPT Scores	-.792	-.156	-.084				.431
Empathy Quotient	.665	.177	.071				.508
Systemizing Quotient	-.443	-.073	-.046				.659

Note. CFMT+ = Cambridge Face Memory Test – Long Version, CCMT = Cambridge Car Memory Test, CFPT = Cambridge Face Perception Test, CCPT = Cambridge Car Perception Test, EQ = Empathy Quotient, SQ = Systemizing Quotient.

** $p < .05$.*

2.3. Summary

Correlational analyses revealed that birthweight and percentile (a combination measure of birthweight and gestation) were related to face memory scores and face perception scores, respectively. Regression analyses revealed that percentile predicted only face perception scores. The findings reported here are consistent with those observed in previous group-based reports in children that show perception-selective face deficits remain stable throughout childhood (i.e. Perez- Roche et al., 2017), suggesting that perinatal effects on face

recognition ability are face-selective and may persist over time. However, these findings do not rule out the possibility that individuals may also show processing differences in the face processing strategy itself. Given that all participants scored within-normal range on the standardized tasks of face recognition ability, accuracy scores as a screening tool alone may mask the use of successful (but atypical) compensatory strategies in individuals with sub-clinical face recognition deficits. To address this possibility, a second experiment was conducted.

3. Experiment 2

To investigate whether perinatal experiences (i.e. birth weight, gestation, and percentile) influence the face-processing strategy itself, a subset of 51 participants from experiment 1 were invited into Bournemouth University to free-view a set of static images displaying social scenes while their eye movements were recorded.

3.1. Method

3.1.1. Participants

A subset of 51 adults who took part in Experiment 1, also took part in this experiment. Adults who had problems with their vision (i.e. weak eyes and short-sightedness, $n = 2$) or had a diagnosed or suspected developmental or neurological disorder ($n = 7$) were excluded from analysis. This resulted in a total of 42 adults (34 = females, 39 = right-handed), aged 18- 47 years (M = 22.74, SD= 5.50), in the final analysis. Reported birth weight ranged from 1kg to 4.65kg (M = 3.17, SD = 0.69): 35 participants were classed as a typical birth weight (>2.5kg; M = 3.46, SD = 0.41), 6 participants were classed as a low birth

weight (<2.5kg; *M*= 2.14, *SD* = 0.38), and 1 participant was classed as a very low birth weight (<1.5kg). Gestation period ranged from 25 weeks to 47 weeks (*M* = 39.19, *SD* = 3.30). Percentiles ranged from the 0^{th} to 99^{th} percentile (*M* = 43.29, *SD* = 30.92); 9 participants were below the 10^{th} percentile (*M* = 3.0, *SD* = 3.17) and experienced a low weight for gestational age. Participants provided written consent and were either compensated with course credit or a £5 Amazon voucher for their time.

3.1.2. Apparatus and materials

"Social scenes" paradigm. Participants free-viewed a "social scenes" task (Bobak et al., 2017). We felt this paradigm was suitable for two reasons. First, it has been shown to illuminate reliable differences in scanning patterns between typical perceivers, super recognizers, and individuals with clinical face recognition deficits (e.g. Bobak et al., 2017). Second, the faces are presented within their natural context, which is thought to be more fruitful when analysing featural fixations (Birmingham, Bischof, & Kingstone, 2008). Specifically, the stimuli in this task includes 20 experimental images and 5 filler images (all presented in colour). The experimental images display the faces and bodies of people engaged in various social activities, such as a family having a picnic or a group of friends eating a meal. The images always include between two and six individuals, who naturally engage with each other (i.e. they do not face the camera). Conversely, the filler images depict natural scenes that do not contain people (e.g. a woodland) and are included to keep participants naïve to the aims of the experiment. Participants view all 25 images in a random order, with an

exposure time of 5s per image, and a centrally positioned fixation cross prior to each stimulus presentation (to confirm retinal attention). Running the task takes approximately five minutes. Example stimuli are displayed in *Figure 4*.

All images were adjusted to 27.09cm in length and 18.07cm in height and subtended 25.44 and 17.13 degrees of visual angles, respectively, when viewed from a distance of 60cm. Eye movements were recorded using the Eyelink 1000 system (SR Research Ltd, Canada) and eye position was monitored through an infrared CCD video camera that was placed on the table in front of participants. As no participants reported problems with their vision, the right eye was always recorded. A chinrest with headrest was used to stabilise participant head position. Eye movements were analysed using Eyelink Data Viewer software (SR Research Ltd). Three levels of areas of interest (AOIs) were pre-drawn onto the 20 experimental images (and analyses were not performed on the 5 filler images). The first level contains the bodies of each individual (taken from below the chin), and the faces of each individual (including hair and ears). In the next level, face region is divided into two separate AOIs; the inner facial features (eyes, nose, mouth and the spaces immediately between them) and the hair (outer facial features including hair and ears). The “inner facial features” are further subdivided into specific feature AOIs, for the eyes, nose, and mouth. We investigated the percentage of dwell time allocated on average to each AOI across the 20 experimental trials.

Figure 4. Example stimuli for Experiment 2, taken from Bobak et al. (2016). Black lines represent areas of interest (AOIs). All images were displayed in colour.

3.1.3. Design

There were three (continuous) independent variables: birth weight, gestation, and percentile. There were nine AOIs, so there were nine dependent variables in total: body dwell time; face dwell time; hair dwell time; inner features dwell time; eyes dwell time; left eye dwell time, right eye dwell time; nose dwell time; and mouth dwell time. We included overall eye dwell time and the dwell time for each eye separately as previous research has investigated both the eyes as a single region (e.g. Peter & Eckstein, 2012) and as two distinct regions (e.g. Vinette et al., 2004). Analysis reflects the fact that some of the DVs are not mutually exclusive (i.e. "inner features dwell time" was calculated to include the overall dwell time for the eyes, nose, and mouth); that is, AOIs

are never ‘doubled-up’ within a statistical test.

3.1.4. Procedure

Participants were seated in a quiet room and were informed they were going to view a set of images, and that they should pay attention to each image but allow their eyes to naturally explore the stimuli. Participants were not required to make any responses during the experiment and their eye-movements were recorded for its entire duration. Once participants understood the task demands, they were asked to place their head within the chin rest (to minimise head movements). A 9-point calibration was conducted prior to the experiment. The calibration procedure began with the presentation of a black cross in the centre of a white computer screen; the cross then moved consecutively around the screen until a “good” corneal lock was achieved in each position. This was immediately repeated to validate the calibration. Once the calibration phase was successfully completed, participants immediately began the experiment. As the task was administered in one continuous block, recalibration was not required unless the participant moved or error in fixation prior to image onset indicated recalibration was necessary.

3.2. Results

An outlier check was carried out for each participant. We defined outliers as scores that fall beyond 2 SDs above or below the mean. This method of outlier removal can be used to remove extreme responses within the dataset. We felt extremely short dwell times (i.e. -2SDs below the mean) might represent movement between AOIs, rather than attention to a specific AOI(s). Similarly, we felt extremely long dwell times (i.e. +2SDs above the mean)

might represent a lack of effort and/or motivation, fatigue, or the fact that minimal instructions were given on this free-viewing social scenes task. We chose to remove data based on the objective 2SD criterion because determining the interval over the mean plus or minus two standard deviations is a conventionally used criterion, to determine extreme scores, among researchers (e.g. Mestry et al., 2017; Dalrymple et al., 2017; Dalrymple et al., 2014). Presently, we removed 11 outliers, which came from 10 participants. This meant that 2.91% of data points were removed (42 x 9 = 378 data points; 11 outliers/378 *100 = 2.91). Statistical analyses were performed on the remaining 32 participants. Despite being commonplace, this method of outlier removal has been criticised for being problematic, as both the mean and SD are influenced by outliers (Leys, Ley, Klein, Bernard, & Licata, 2013). Given this, we provide a supplementary analysis of our results without any data removal (see *Appendix B*). Although data removal did not drastically affect our results and associated interpretations, all key differences are highlighted to the reader in *Appendix B*. Accordingly, our chosen method of outlier removal is revisited within the discussion with a focus on its disadvantages and some alternative approaches. As our results differed depending on the inclusion or exclusion of outliers, we felt the presentation of both analyses (either in the thesis itself or as supplementary information in *Appendix B*) was appropriate.

A Shapiro-Wilk test was used to assess the sample distribution of all variables. Percentile (W (32) = .951, p = .158), dwell time on the body (W (32) = .966, p = .390), dwell time on the face (W (32) = .967, p =.429), dwell time on the inner features overall (eyes, nose, and mouth) (W (32) = .981, p =.840),

including the nose (W (32) = .950, p =.146) and eyes (W (32) = .954, p = .189), left eye specifically (W (32) = .943, p =.091), and dwell time on the hair (W (32) = .975, p = .653) were normally distributed. Birth weight (W (32) = .918, p = .018), gestation (W (32) = .878, p = .002), right eye specifically (W (32) = .912, p =.012) and dwell time to the mouth (W (32) = .924, p = .026) deviated significantly from normality. Transformations did not normalise the data, so non-parametric statistics are reported where appropriate.

Correlational analyses were used to examine the relationship between participants' birth weight, gestation period, percentile, and dwell time (%) on bodies, faces, hair, inner features, eyes (overall and individual), nose, and mouth (see *Table 6*). Two-tailed correlations revealed moderately positive relationships between birth weight and dwell time on the hair (r_s = .356, p = .023), and percentile and dwell time on the hair (r_p = .327, p = .034). A moderately negative relationship was observed for percentile and dwell time on the eyes (r_p = -.359, p = .022) and for percentile and right eye dwell time (r_s = -.414, p = .018). A weak negative relationship between percentile and dwell time on the inner features overall was also approaching significance (r_p = -.289, p = .054). Scatterplots were produced only for these relationships (see *Figure 5*). No other correlations were significant (all ps> .115). Higher birth weights and percentiles were associated with greater dwell times on the hair, reduced dwell times on the eyes (overall and right eye in particular) and the inner features overall.

Table 6. Correlation coefficients of birth weight, gestation, and percentile with dwell time (%) on bodies, hair, faces, and inner features, including the eyes (overall and individual), nose and mouth. p-values are reported in parentheses and significant correlations are highlighted in bold.

		Hypothesised Predictors	
	Birth weight	**Gestation**	**Percentile**
Body Dwell Time	-.086 *(.320)*	-.173 *(.172)*	.036 *(.423)*[a]
Face Dwell Time	-.036 *(.422)*	.037 *(.420)*	-.034 *(.426)*[a]
Hair Dwell Time	**.356 *(.023)***	-.141 *(.221)*	**.327 *(.034)*[a]**
Inner Features Dwell Time	-.218 *(.115)*	.070 *(.351)*	-.289 *(.054)*[a]
Eyes Dwell Time	-.192 *(.146)*	.051 *(.391)*	**-.359 *(.022)*[a]**
Nose Dwell Time	-.074 *(.344)*	-.034 *(.426)*	.023 *(.451)*[a]
Mouth Dwell Time	.177 *(.166)*	-.021 *(.454)*	.181 *(.161)*
Left Eye Dwell Time	.051 *(.781)*	-.057 *(.755)*	.000 *(.999)*[a]
Right Eye Dwell Time	-.295 *(.101)*	.052 *(.779)*	**-.414 *(.018)***

[a] *Parametric coefficients reported.*

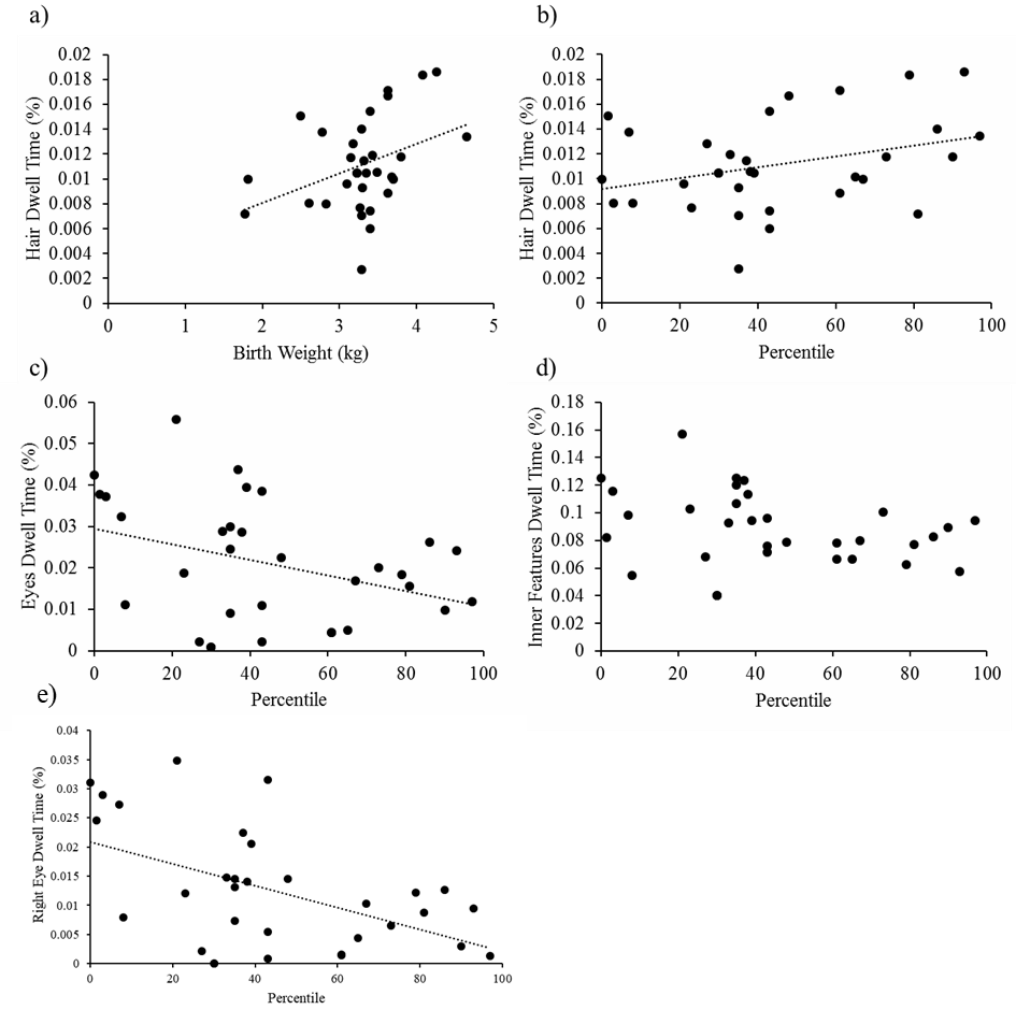

Figure 5. Scatterplots to show the relationship between a) birth weight and hair dwell time; b) percentile and hair dwell time; c) percentile and eyes dwell time; d) percentile and inner features (overall) dwell time; and e) percentile and right eye dwell time. Coefficients signified only moderately positive (a and b), moderately negative (c and e), or weak negative (d) relationships. The trendline was not applied for (d), where there was only a trend towards significance.

As expected, all three infancy-related hypothesised predictors were correlated with each other (birthweight and gestation, r = .126; birthweight and percentile, r = .683; percentile and gestation, r = -.271). Likewise, participants'

scores across the various dwell time dependent measures were correlated with each other (see *Table 7* for correlation matrix). Although these associations varied in strength, from weak to strong, the following regression analyses reflect the fact that some of the DVs are not mutually exclusive. In other words, conducting separate regression analyses removed potentially confounding contributions. As such, multicollinearity is not a problem.

Table 7. Correlation matrix of the dwell time dependent variables. Correlations ranged from weak to strong.

	Dependent Measures								
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Body Dwell Time (1)	1	-.729	.173	-.736	-.610	-.610	.111	-.670	-.516
Face Dwell Time (2)	-	1	-.292	.838	.406	.840	.107	.555	.314
Hair Dwell Time (3)	-	-	1	-.256	.157	-.218	-.423	.176	.110
Inner Features Dwell Time (4)	-	-	-	1	.684	.799	-.138	.586	.644
Eyes Dwell Time (5)	-	-	-	-	1	.422	-.714	.837	.951
Nose Dwell Time (6)	-	-	-	-	-	1	-.116	.503	.364
Mouth Dwell Time (7)	-	-	-	-	-	-	1	-.548	-.719
Left Eye Dwell Time (8)	-	-	-	-	-	-	-	1	.673
Right Eye Dwell Time (9)	-	-	-	-	-	-	-	-	1

Note. Some DVs are not mutually exclusive (e.g. "inner features dwell time was calculated to include the overall dwell time for the eyes, nose, and mouth).

Multiple linear regressions

Separate multiple linear regressions were conducted to predict dwell times from each infancy-related IV: birthweight, gestation, and percentile. As some of the DVs are not mutually exclusive (i.e. "inner features dwell time" is calculated to include the overall dwell time for the eyes, nose, and mouth), any DVs that related to each other in this way were further divided and also entered into separate regression models. This resulted in a total of six regression models. In each model, there is one continuous IV (i.e. birthweight, gestation, or percentile) and multiple DVs. Body dwell time, hair dwell time (outer facial features including the hair and ears), and inner features dwell time were entered into the first 'set' of regressions. Left eye dwell time, right eye dwell time, nose dwell time, and mouth dwell time were entered into the second 'set' of regressions. This second 'set' was conducted in order to separate the inner facial features. As we are simply quantifying the relationships between variables, the infancy-related IVs were each separately entered into SPSS as a DV and the dwell time measures were entered into SPSS as predictors. The following results are reported as entered into the model (i.e. dwell time measures predict infancy-related measures) but are interpreted as the reverse (i.e. infancy-related measures predict dwell time measures).

Birthweight

Table 8 provides a summary for the birthweight regression analysis predicting birthweight from body dwell time, hair dwell time, and inner features dwell time. The model was non-significant; $F(3, 28) = 2.442$, $p = .085$,

accounting for 12.2% of the variance (Adj. R^2 = .122). As presented in *Table 8*, all predictors were non-significant (all *p*s ≥ .070), although hair dwell time indicated a trend towards significance (β = .334, *t* = 1.885, *p* = .070). Birthweight did not predict body dwell time, hair dwell time, or inner features dwell time measures.

Table 8. Summary of the multiple linear regression analysis predicting birthweight from body dwell time, hair dwell time, and inner features dwell time.

Variable	*t*	b	β	*F*	R^2	Adj. R^2	*p*
Overall Model				2.442	.207	.122	.085
Body Dwell Time	-1.404	-15.785	-.395				.171
Hair Dwell Time	1.885	52.799	.334				.070
Inner Features Dwell Time	-1.493	-9.510	-.429				.147

Table 9 provides a summary for the birthweight regression analysis predicting birthweight from left eye dwell time, right eye dwell time, nose dwell time, and mouth dwell time. The model was significant; *F*(4, 27) = 2.801, *p* = .046, accounting for 18.9% of the variance (Adj. R^2 = .189). As presented in *Table 9*, right eye dwell time was a significant predictor of birthweight (β = -.693, *t* = -2.957, *p* = .006); all other predictors were non-significant (all *p*s ≥ .085). Birthweight significantly predicted right eye dwell time measures but was unrelated to left eye dwell time, nose dwell time, and mouth dwell time measures.

Table 9. Summary of the multiple linear regression analysis predicting birthweight from left eye dwell time, right eye dwell time, nose dwell time, and mouth dwell time.

Variable	*t*	b	β	*F*	R^2	Adj. R^2	*p*
Overall Model				2.801	.293	.189	.046*
Left Eye Dwell Time	1.788	39.067	.407				.085
Right Eye Dwell Time	-2.957	-39.545	-.693				.006*
Nose Dwell Time	-.095	-.689	-.018				.925
Mouth Dwell Time	-.208	-1.777	-.048				.837

** $p < .05$.*

Gestation Period

Table 10 provides a summary for the gestation period regression analysis predicting gestation period from body dwell time, hair dwell time, and inner features dwell time. The model was non-significant; $F(3, 28) = .171$, $p = .915$, accounting for 9.1% of the variance (Adj. $R^2 = -.018$). As presented in *Table 10*, all predictors were non-significant (all *p*s ≥ .641). Gestation period did not predict body dwell time, hair dwell time, or inner features dwell time measures.

Table 10. Summary of the multiple linear regression analysis predicting gestation period from body dwell time, hair dwell time, and inner features dwell time.

Variable	*t*	b	β	*F*	R^2	Adj. R^2	*p*
Overall Model				.171	.018	-.087	.915
Body Dwell Time	-.471	-23.190	-.148				.641
Hair Dwell Time	.178	21.801	.035				.860
Inner Features Dwell Time	-.041	-1.144	-.013				.968

Table 11 provides a summary for the gestation period regression analysis predicting gestation period from left eye dwell time, right eye dwell time, nose dwell time, and mouth dwell time. The model was non-significant; $F(4, 27) = .355$, $p = .838$, accounting for 9.1% of the variance (Adj. $R^2 = -.091$). As presented in *Table 11,* all predictors were non-significant (all *p*s ≥ .253). Gestation period did not predict left eye dwell time, right eye dwell time, nose dwell time, or mouth dwell time measures.

Table 11. Summary of the multiple linear regression analysis predicting gestation period from left eye dwell time, right eye dwell time, nose dwell time, and mouth dwell time.

Variable	*t*	b	β	*F*	R^2	Adj. R^2	*p*
Overall Model				.355	.050	-.091	.838
Left Eye Dwell Time	-.363	-36.667	-.096				.720
Right Eye Dwell Time	1.169	71.165	.318				.253
Nose Dwell Time	-.231	-7.609	-.049				.819
Mouth Dwell Time	.466	18.114	.125				.645

Percentile (combination measure of birthweight and gestation period)

Table 12 provides a summary for the percentile regression analysis predicting percentile from body dwell time, hair dwell time, and inner features dwell time. The model was significant; $F(3, 28) = 3.097$, $p = .043$, accounting for 16.9% of the variance (Adj. $R^2 = .169$). As presented in *Table 12*, inner features dwell time was a significant predictor of percentile ($\beta = -.641$, $t = -2.289$, $p = .030$) and body dwell time indicated a trend towards significance ($\beta = -.536$, $t = 1.471$, $p = .060$); hair dwell time was non-significant ($p = .152$). Percentile significantly predicted inner features dwell time and appears to be related to body dwell time but was unrelated to hair dwell time.

Table 12. Summary of the multiple linear regression analysis predicting percentile from body dwell time, hair dwell time, and inner features dwell time.

Variable	*t*	b	β	*F*	R^2	Adj. R^2	*p*
Overall Model				3.097	.249	.169	.043*
Body Dwell Time	-1.960	-1026.19	-.536				.060
Hair Dwell Time	1.471	1919.54	.253				.152
Inner Features Dwell Time	-2.289	-679.07	-.641				.030*

** $p < .05$.*

Table 13 provides a summary for the percentile regression analysis predicting percentile from left eye dwell time, right eye dwell time, nose dwell time, and mouth dwell time. The model was significant; $F(4, 27) = 5.103$, $p = .003$, accounting for 34.6% of the variance (Adj. $R^2 = .346$). As presented in *Table 13*, right eye dwell time was a significant predictor of percentile ($\beta = -.908$, $t = -4.313$, $p < .001$) and left eye dwell time indicated a trend towards significance ($\beta = .381$, $t = 1.864$, $p = .073$); nose dwell time ($p < .739$) and mouth dwell time ($p < .278$) were non-significant. Percentile significantly predicted right eye dwell time and appears to be related to left eye dwell time but was unrelated to nose dwell time and mouth dwell time measures.

Table 13. Summary of the multiple linear regression analysis predicting percentile from left eye dwell time, right eye dwell time, nose dwell time, and mouth dwell time.

Variable	*t*	b	β	*F*	R^2	Adj. R^2	*p*
Overall Model				5.103	.431	.346	.003*
Left Eye Dwell Time	1.864	1749.79	.381				.073
Right Eye Dwell Time	-4.313	-2477.79	-.908				.000**
Nose Dwell Time	.337	104.591	.056				.739
Mouth Dwell Time	-1.106	-405.839	-.229				.278

** p < .05, **p < .001.*

Disentangling the contribution of birthweight and sex

Given that both the birthweight and percentile regression analyses yielded significant results for the inner facial features (in particular, right eye dwell time), it remains unclear what drives this variation. Indeed, it is possible that birthweight predicts variation in visual processing strategies. However, because percentile is calculated using birthweight, gestation, and sex (i.e. male or female), we cannot rule out the possibility that sex may instead (or at least in-part, also) drive this variation. To address this possibility, we present birthweight regression analyses separately for both males and females.

Males

Table 14 provides a summary for the birthweight regression analysis predicting birthweight from body dwell time, hair dwell time, and inner features dwell time for males only (n = 7). The model was non-significant; $F(4, 2)$ =

6.401, p = .140, accounting for 78.3% of the variance (Adj. R^2 = .783). As presented in *Table 14*, all predictors were non-significant (all ps ≥ .101). In the male sample, birthweight did not predict body dwell time, hair dwell time, or inner features dwell time measures.

Table 14. Summary of the multiple linear regression analysis predicting birthweight from left eye dwell time, right eye dwell time, nose dwell time, and mouth dwell time for males only.

Variable	*t*	b	β	*F*	R^2	Adj. R^2	*p*
Overall Model				6.401	.928	.783	.140
Left Eye Dwell Time	-.098	-3.306	-.021				.931
Right Eye Dwell Time	-2.895	-75.168	-.948				.101
Nose Dwell Time	-1.675	-20.088	-.420				.239
Mouth Dwell Time	.412	10.851	.159				.720

Females

Table 15 provides a summary for the birthweight regression analysis predicting birthweight from body dwell time, hair dwell time, and inner features dwell time for females only (n = 25). The model was non-significant; F(4, 20) = 2.084, p = .121, accounting for 15.3% of the variance (Adj. R^2 = .153). As presented in *Table 15*, left eye dwell time and right eye dwell times both indicated a trend towards significance (β = .582, t = 1.975, p =.062 and β = -.590, t = -2.058, p =.053, respectively); nose dwell time (p < .450) and mouth dwell time (p < .470) were non-significant. In the female sample, birthweight

appears to be related to left eye and right eye dwell times but unrelated to nose dwell time and mouth dwell time measures.

Table 15. Summary of the multiple linear regression analysis predicting birthweight from left eye dwell time, right eye dwell time, nose dwell time, and mouth dwell time for females only.

Variable	***t***	**b**	**β**	***F***	R^2	Adj. R^2	***p***
Overall Model				2.084	.294	.153	.121
Left Eye Dwell Time	1.975	49.041	.582				.062 *
Right Eye Dwell Time	-2.058	-30.030	-.590				.053 *
Nose Dwell Time	-.771	-6.354	-.169				.450
Mouth Dwell Time	.736	6.884	.206				.470

** p < .07 (a trend towards significance).*

3.3. Summary

Correlational analyses revealed that percentile (a combination of birth weight and gestation) was related to, and predicted, only inner feature (eyes, nose, mouth) dwell times, in particular the eye region. In contrast to our prediction, higher percentiles were associated with shorter dwell times to the inner features, providing no evidence for reduced optimum processing with reduced face perception skill. These results suggest that individuals with sub-clinical deficits do not show atypical face processing strategies and perform within-normal range on standardized tasks of face recognition ability. It is important to note, however, that when outliers are not excluded from analysis,

all results become non-significant (nor are they approaching significance; *see Appendix B*). To disentangle the (relative) contribution of birthweight and sex, we conducted two separate regressions for males and females; only the female data showed a trend towards significance. It is possible sex (i.e. male or female) might drive (at least) some of the variation in face processing strategy. However, given that the current sample sizes are small ($n_{male} = 7$, $n_{female} = 25$), and outlier inclusion affects the current interpretation, this possibility requires additional support from future research. These issues, as well as the combined findings of the explicit and implicit methodologies, are considered more thoroughly in the discussion.

4. Discussion

This study investigated the developmental trajectory of face recognition impairments to determine whether perinatal effects on face recognition ability in childhood merely delay the development of typical face recognition skills or persist into adulthood. Principally, Experiment 1 suggests that adverse perinatal effects on face recognition ability in childhood persist into adulthood. These effects were found to be face-selective and were limited to perceptual processes. Although this relationship was observed, no participants were clinically impaired on behavioural tasks of face recognition ability (Experiment 1) or exhibited atypical processing strategies in the face-processing strategy itself, when free-viewing social scenes (Experiment 2). Although, with increased power, it is possible effects around the inner features, and particularly the eyes, will emerge. These findings are in line with the prediction that individual

differences in face recognition ability that result from adverse perinatal influences are subtle, although a greater number of low birth weight individuals need to be recruited to test this possibility.

4.1. Memory versus perception

Our principal finding builds upon the initial evidence provided by Perez-Roche et al. (2017) that suggests early deficits impair the development and specialization of the face- processing system through to adulthood. Specifically, we found birth weight for gestational age was related to, and predicted, face perception but not face memory scores in an adult sample; the same dissociation reported in Perez-Roche et al.'s (2017) child sample. This finding is demonstrated more strongly in the present study as adults completed dominant tests of both perception and memory independently. The Facial Memory Subtest used by Perez-Roche et al. (2017) is taken from TOMAL (Test of Memory and Learning; Reynolds & Voress, 2007). In this task, participants are required to remember an increasing number of faces over 12 trials; first, participants view the faces before moving on to another page where they are asked to indicated which identities they had just seen amongst a set of distractors. Given that the Facial Memory Subtest used by Perez-Roche et al. (2017) does not engage perceptual processes without a memory component (i.e. participants are always remembering 'just-seen' faces), it is likely these processes were somewhat confounded in the child sample. Although the present study suggests sub-clinical face perception atypicalities persist into adulthood, and Perez-Roche et al.'s (2017) data also suggests that they do not improve over the course of development, it remains possible that face recognition impairments may also

change throughout childhood. The data collected from children as part of this wider project will speak to this point.

Research has indicated a dissociation between face perception and face memory processes in adults, but not children, with DP (Dalrymple, Garrido, & Duchaine, 2014), suggesting that face perception can improve prior to, or during, adulthood. It is important to note, however, that this study was comprised of a very small sample size ($n_{adults} = 16$; $n_{children} = 8$), so the results may have occurred by chance. As such, we cannot rule out the possibility that some individuals may have overcome their face perception (and potentially face memory) atypicalities at some point in development. Partial support for this notion comes from the decrease in prevalence estimates from childhood (4-5%; Bennetts et al., 2017) to adulthood (2-2.5%; Bowles et al., 2009) in DP. However, it is important to note these studies used very different approaches and the age-appropriate face matching tests used may simply overestimate the prevalence of face recognition impairments in the child population. In line with Perez-Roche et al. (2017), the levels of face recognition impairments did not decrease with age in the childhood prevalence sample (Bennetts et al., 2017), suggesting that "catch-up" processes occur some point later in development; at present, fifteen years of age is a "lower bound estimate" of when in development these might occur for sub-clinical atypicalities (i.e. Perez-Roche et al., 2017).

Likewise, face memory was unimpaired in our sample. We did not come to this conclusion because there was an absence of correlation regarding the face memory performance. Instead, this was concluded because none of our

participants scored lower than 52.78% (or 38 raw score). We compared our present CFMT scores with appropriate published norming data. For example, our mean CFMT score (M = 56.22) and standard deviation (SD = 8.02) was comparable to the controls reported in Bate et al.'s (2014) study (M = 59.6; SD = 7.6); conversely, the average score for DPs was 36.1. Similarly, our mean CFMT score and standard deviation was comparable to the controls reported in Duchaine and Nakayama's (2006b) study (M = 57.6; SD = 11); in this study, the average score for DPs was slightly lower, at 28.06. Given that our descriptive statistics were comparable to previously tested samples within the typical population, and were not as low (or lower) than previously tested samples of clinically impaired individuals, we believe that none of our participants were clinically impaired on face memory, as measured by the CFMT. Further, although not empirically tested or (in)formally asked by the experimenter, no participants offered anecdotes or instances where they failed to recognize close friends or family, nor did they report any lifelong difficulties with their face recognition. The absence of an impairment in face memory in our sample raises the intriguing question of whether face memory impairments can, or might, also improve (in some individuals) across development. As our sample size is currently a little small, it is possible differences in face memory may become apparent with increased power, but inspection of the correlation coefficients suggests only very weak relationships between percentile and the face memory task, and effects disappear when entered into the regression model. Research has shown face perception processes to mature earlier than face memory processes (Weigelt et al., 2014), which means that it may simply be more

difficult to detect atypical performance in memory in early development than perception (i.e. matching) tasks where mechanisms are fully developed. The observation that face memory was also unimpaired in Perez-Roche et al.'s (2017) child sample suggests two possibilities. First, it is possible face memory processes are unaffected by perinatal influences. Second, perinatal effects on face recognition ability may be qualitatively different in childhood and adulthood; it is possible age-appropriate and more sensitive measures (i.e. those that don't confound perception and memory processes) of face memory in early and middle childhood will detect atypicalities in memory (beyond what is already expected from individuals of that age) in younger samples. It would also be informative to calculate the proportion of individuals below the 10th percentile (i.e. those that were born small for gestational age) who meet the criteria for DP; if this falls above the standard ~2% prevalence figure (Bowles et al., 2009), it might be able to be considered a risk factor. It is important to note that this prevalence figure has been criticised of simply being a statistical artefact, as the procedure typically identifies the bottom 2% of the population in a normally distributed sample (Bate & Tree, 2018). No individuals met the criteria for DP in our sample and we cannot directly compare Perez-Roche et al.'s (2017) group-based approach with our correlational design, so this a matter for future research. Critically, if some children "outgrow" their face perception atypicalities, we would expect a subset of Perez- Roche et al.'s (2017) small for gestational age children to show improved scores on perception tasks later in life. Longitudinal work is needed to test this possibility, and studies of individual cases may also help to determine what proportion of individuals

"catch-up" or "outgrow" their face recognition deficits and what characteristics, if any, differ between individuals who continue to struggle with face recognition in adulthood, and those that do not.

Preterm severity may be one characteristic that distinguishes between developmental delays or permanent deficits. A longitudinal cohort study (that previously assessed children at 5 years of age) found that children born extremely preterm (23-27 weeks) were at increased risk of long term cognitive deficits when re-assessed at age 18, whereas children born only moderately (32-36 weeks) or very (28-31) preterm functioned like their term-born counterparts at 18 years of age (Lundequist et al., 2014). Although this study did not directly test face recognition ability, it raises the possibility that if a "catch-up" function emerges is related to the severity of prematurity. Specifically, shorter gestation periods may predict greater developmental delays, or (in the shortest cases) permanent deficits. Only seven adults were classified (UNICEF & WHO, 2004) as premature in the present study; four individuals were moderately preterm, two were very preterm, and one was extremely preterm. Consequently, we did not have enough power to detect group differences.

Further, although the precise relationship between adverse perinatal experiences and atypical outcomes is unknown, it is plausible there will be linear differences between moderately, very, and extremely preterm individuals. Evidence from a variety of sources supports this suggestion. For example, one prospective longitudinal study found that although term and pre-term individuals significantly differed in somatic, neuro-motor, cognitive, and socio-emotional developmental domains, a subgroup of pre-term children without

overt school difficulties were characterized by less adverse perinatal experiences and better mental and motor development (Baarn Ultee, Gunning, Soepatmi, & de Leeuw, 2006). A more recent study of an Australian cohort found that adverse perinatal experiences (i.e. prematurity, low birth weight, and being small for gestational age) were negatively associated with cognitive outcomes in childhood (Chen et al., 2014). In this study, very low birth weight children scored -0.55 SDs lower on tests assessing cognitive school readiness whereas moderately low birthweight children only scored -0.23 SDs lower than controls. Likewise, neuroimaging studies have found correlations between regional brain volumes (in the sensorimotor and mid-temporal cortices), cognitive measures, and perinatal variables among preterm children, at 8 years of age (Peterson et al., 2000). This research suggests an indirect mechanism by which adverse perinatal experiences may impact cognitive outcomes. Specifically, prematurity might result in long-term reductions in brain morphology, which in turn is associated with atypical developmental and cognitive outcomes. Future research is needed, however, to investigate whether such findings are the result of obstetric (in-utero) or neonatal (after birth) complications, or some combination of these factors. Nevertheless, if convincing evidence of perinatal influences on later face recognition can be collected and understood, this may potentially be a good focus for early intervention that addresses the development of face processing skills. It is also possible the severity of prematurity differentially affects perception or memory processes. All these possibilities should be the subject of future research and will begin to be addressed as part of our on-going research project as we

continue collect data from children at various stages of development.

Critically, our behavioural data provides novel insights that strengthen the distinction between face perception and face memory; specifically, we, along with Perez-Roche et al. (2017), provide data that suggests the existence of atypical face perception skills across development. Given that face memory, by definition, is impaired in DP, and the current results suggest intact face memory, it is unclear why selective atypical face perception skills have been observed in these samples. It is likely research into dissociations from a clinical population (i.e. individuals with DP) do not readily transfer to the typical population (i.e. individuals with sub-clinical deficits), so one would not expect to see the same pattern of responding between DPs and individuals who have experienced low birth weight for gestational age. One possibility is that atypical face perception skills are not as damaging or pervasive as face memory deficits, and it is likely prior research has failed to detect individuals who only have atypical face perception skills, as people will be unlikely to self- refer. It is also possible atypical face perception skills may be the underpinning cause of sub-clinical impairments. This finding presents important implications for developmental theories of face processing, as sequential cognitive models of face perception suggest intact face perception skills are necessary for successful recognition (e.g. Bruce & Young, 1986).

An alternative, though not mutually exclusive, possibility is that face memory and face perception processes are differentially affected by face familiarity. Although familiar and unfamiliar face processing share some characteristics, such as ‘holistic’ processing (i.e. upright faces are processed as a

unified whole, e.g. Collishaw & Hole, 2000), there is evidence that they also rely on qualitatively different types of information (e.g. Hole, 1994), and are underpinned by distinct neural representations (Natu & O'Toole, 2011). Current theoretical models of face learning suggest that unfamiliar faces are not faces, but they become faces when observers have accumulated sufficient visual experience with an identity (e.g. Megreya & Burton, 2006). As such, less stable (unfamiliar) face representations pose greater challenges to face perception and memory systems than robust (familiar) face representations, due to the additional neural effort and computation required to learn new faces (Natu & O'Toole, 2011). Given that we observed selective deficits in face perception, assessed by a standardized task comprised of unfamiliar faces, it is possible that the fine-grained perceptual discrimination of faces in individuals with sub-clinical deficits is only accrued over time as characters become increasingly familiar to them. Critically, if perinatal effects of face recognition ability are limited to deficits in the perceptual discrimination of unfamiliar faces, we would expect all our sample to show improved scores on perception tasks of familiar faces.

Importantly, if deficits are restricted to unfamiliar faces, it follows that they are relatively temporary. That is, once an individual accumulates sufficient visual expertise with a character and the face is familiar to them, they can discriminate subtle changes within that face. It is also unclear what defines "sufficient" visual experience (i.e. at what point does a face stop being unfamiliar); if perception processes are not immediately functional, this process may be delayed. If this is the case, impairments reflect delays in some aspects of

face learning that “catch-up” as a function of familiarity. It is also intuitive that individuals with sub-clinical deficits do not report difficulties in their day-to-day lives, if their deficits only relate to unfamiliar face processing (i.e. there are less opportunities for social feedback if an unfamiliar character is not recognised optimally). Evidence also suggests that processing may vary for different types of “familiar” faces (specifically, one’s own face, personally familiar faces, and celebrities, e.g. Ma & Han, 2010). As familiar face processing is likely to be confounded with perceptual and (theoretically) conceptual information (Carbon, 2008), as well as ceiling effects, future work should test these possibilities through training paradigms that manipulate familiarity. In sum, these predictions support the notion that perception deficits alone are not as damaging as memory deficits, particularly if they are restricted to the processing of unfamiliar faces.

4.2. Faces versus objects

We also found that adverse perinatal effects on face perception ability were face- selective, as (relative) atypicalities were not observed in a non-face object category (i.e. cars), suggesting that face and object processes rely to a large extent on independent mechanisms in adulthood (e.g. Dennett et al., 2012). As Perez-Roche et al.’s (2017) assessment battery did not include non-face object recognition tests, it is still unclear at what point in development sub-clinical atypicalities might become face-selective. In the typical population, behavioural studies suggest five years of age is an “upper bound estimate” of when in development face and object separation occurs, as face and object processes can be dissociated from this age (Bennetts et al., 2017). However, it is

possible subclinical face-selective atypicalities will only emerge with age once face processing mechanisms are fully developed; that is, children do not exhibit 'impaired' face recognition until their counterparts also show adult-like face processing abilities. Specifically, if face recognition follows a protracted period of development, independent of object recognition, we would expect to see face-selective impairments to emerge at around ten years of age (e.g. Carey & Diamond, 1994; de Heering, Rossion, & Maurer, 2012). Conversely, if face recognition is mature at a very young age, and all subsequent development is a factor of general cognitive development (e.g. Crookes & McKone, 2009; Want, Pascalis, Coleman, & Blades, 2003), we would expect to see face-selective impairments emerge much earlier in development. Many studies are consistent with the conclusion that face-selective processing is mature early in development (e.g. Bennetts et al., 2017; McKone, Crookes, Jefferey, & Dilks, 2012), supporting the hypothesis that face and object processes diverge early in development.

It has also been suggested that face memory and perception undergo separate developmental trajectories, and it is possible these trajectories interact differently across object classes (i.e. face versus non-face objects). Specifically, face perception appears to mature in early childhood and at the same rate as perception for other classes of objects, whereas face memory follows a protracted period of development, until mid-adolescence, that diverges from other classes of objects (Weigelt et al., 2014). It is possible that low birth weight for gestational age has an immediate effect on both faces and objects, but by adulthood it is just restricted to faces. If object deficits co-present earlier in

development, these may also be limited to perception processes. However, research has also shown face memory and object memory to have similar developmental trajectories (Bennetts et al., 2017). It is unclear exactly why these results differ as there were several methodological differences between the two studies, however, this does highlight the importance of designing theoretically appropriate methods to assess face recognition ability in children. Knowing at what age face and object separation typically occurs would also allow healthcare providers to assess, monitor, and provide interventions to children at risk of failing to separate these processes.

4.3. Alternative methodologies

To reiterate, our results suggest adverse perinatal effects on face recognition ability persist into adulthood, however, the underpinnings of this effect remain unclear. It is unlikely the impairments reported in our sample are accounted for by general perceptual processing mechanisms (i.e. deficits were face-selective) or co-occurring socio-developmental disorders (i.e. individuals with a diagnosed or suspected developmental disorder were excluded from analysis). It is also a possibility that adverse perinatal experiences impair social functioning, which in turn affects face processing ability itself. In addition, results from Experiment 2 did not find evidence for the use of atypical processing strategies, suggesting that differences in allocation of visual attention also do not underpin sub-clinical face recognition deficits. In contrast to our prediction, results showed that higher percentiles were associated with shorter dwell times to the inner features (and in particular the eyes), suggesting that perinatal factors do not relate to processing in terms of dwell time. It is unusual

we didn't observe the typical finding for increased attention to the eyes (as indexed by higher dwell time percentages; e.g. Peterson & Eckstein, 2012), given that we tested individuals from the typical population. Although we assumed increased dwell times reflect the allocation of attentional resources and an optimal processing strategy (Hills & Willis, 2016), it is possible the patterns we observed instead simply reflect individual differences in the speed or efficiency with which individuals' process faces. Given that higher percentiles were associated with shorter dwell times, it is plausible individuals who are appropriate for gestational age extract and accumulate facial information quickly and efficiently. Conversely, increased dwell times might reflect a successfully applied compensatory strategy. Specifically, individuals with lower percentiles might focus their attention to diagnostic facial features, such as the eyes, but are simply less efficient at extracting this information. Evidence from a variety of sources supports this hypothesis. Within the developmental literature, the mean fixation duration of individuals with dyslexia was significantly shorter after receiving training to improve their reading skills, suggesting that training improved the speed of extracting visual information (Judica, De Luca, Spinelli, & Zoccolotti, 2002). Similarly, short-looking infants often perform better than long-looking infants on perceptual cognitive tasks (e.g. Colombo, Mitchell, Coldren, & Freeseman, 1991; Coombo & Mitchell, 1990; Cooper et al., 1990), suggesting that short-lookers simply process stimuli more rapidly than long-lookers. It is possible that short-lookers' advantage in processing speed may be the result of structural factors. Specifically, the structural integrity of the white matter tracts associated with the superior

longitudinal fasciculus are thought to make a prominent contribution to cognitive processing speed (Turken et al., 2008). This underlying mechanism may be what drives atypicalities that present in low for gestational weight adults, and requires attention in future research. Relatedly, previous research supports the hypothesis that the female advantage often observed in facial expression recognition is associated with greater female attention to the eyes (Hall, Hutton, & Morgan, 2008). Presently, we observed a negative right eye dwell time association and a positive left eye dwell time association. However, when further analyses were conducted for males and females separately, the left eye dwell time association only remained positive for the females, suggesting that females (and not males) were driving the effect of increased attention to the left eye. However, as the final analysis was conducted on a small sample size (N = 32), the effects we observed are underpowered and are likely to be noisy. Alternatively, it is possible that individuals with reduced face perception skills do use atypical processing strategies, but the social scenes task is simply not sensitive enough to differentiate atypical skills in the typical population.

The differences between the stimuli of the face perception task and the social scenes paradigm may be able to account for our lack of convergence across experiments 1 and 2. In the face perception task (i.e. CFPT; Bowles et al., 2009), images are cropped to an oval to show only the internal facial features, are presented in isolation (i.e. against a plain background), and are greyscale. In contrast, the social scenes paradigm (Bobak et al., 2016) presents faces within their natural context, which includes multiple persons, backgrounds, and colour. Although presenting faces in this way is thought to be

more informative when analysing featural fixations (e.g. Birmingham, Bischof, & Kingstone, 2008) and was one of our reasons for choosing this task, it is possible that because the focus was not on faces and participants were only instructed to naturally explore the stimuli (i.e. they did not have to carry out a task, and instructions were minimal), the task does not differentiate individuals with sub-clinical deficits.

Again, if perception processes are impaired as a function of facial familiarity in individuals with sub-clinical deficits, our finding of shorter dwell times on the inner features may be better explained by evidence that familiar and unfamiliar face processing rely on qualitatively different types of information (e.g. Hole, 1994). Whilst internal facial features (e.g. eyes, nose, mouth) are critical for familiar face recognition, external features (i.e. ears and hairstyle) are used more frequently for unfamiliar face processing (Young et al., 1986). Our eye-movement data is consistent with the fact participants were unfamiliar with the characters in the social scenes and does not necessarily indicate the use of atypical processing strategies. Further, if perception impairments improve over time with increasing familiarity, it is likely we would observe optimal internal featural fixations for familiar faces; it is possible sub-clinical perceptual impairments simply delay this process. Alternatively, if perception deficits persist across face 'levels' (i.e. familiar versus unfamiliar faces), we may also see disruptions to regions critical for face identification (such as the eyes or the nose; Taylor et al., 2001 and Peterson & Eckstein, 2012, respectively) that may be quantitative in nature. It is unclear how individuals with sub-clinical deficits process both familiar and unfamiliar faces, and how

this might change over time. Given that our findings are based in differences in face perception, it would be informative for future research to use eye-movement recording techniques in the context of perception-based matching or judgement tasks.

Our findings also highlight the value of research that integrates behavioural and implicit (i.e. eye-tracking) methodologies. Given that implicit measures absolve issues relating to task difficulty and there are considerable difficulties hindering the early detection of atypical face recognition difficulties, eye-tracking may be best used in the context of childhood studies, as the tests are shorter and require much less input from participants (Turati et al., 2010). Atypical patterns of eye-movements have been documented in typically developing children when processing other-age faces (Hills & Willis, 2016). It is also possible atypical face recognition fixation patterns may be qualitatively different in childhood and adulthood and may simply reflect the delayed development of effective processing strategies in individuals who have experienced a low birth weight for gestational age. If this is the case, we would expect to reliably see the use of atypical strategies in younger children, but a move to more optimal mechanisms prior to, or during, adulthood.

4.4. Potential applications

Relatedly, if convincing evidence of perinatal influences on the use of atypical processing strategies can be collected and understood, instructing eye movements may be a suitable target for the rehabilitation of individuals with atypical face recognition abilities, although effects are likely to be small. Perceptual training paradigms have shown promising results in child (Brunsdon,

Coltheart, Nickels, & Joy, 2006; Schmalzl, Palermo, Green, Brunsdon, & Coltheart, 2008), adolescent (Bate et al., 2014) and adult (DeGutis, Bentin, Robertson, & D'Esposito, 2007; DeGutis, Cohan, & Nakayama, 2014) cases of DP. Although these training programmes differ in several methodological ways and do not solely focus on instructing the allocation of visual attention, they highlight two key theoretical issues. First, rehabilitative progress has only been achieved in perceptual aspects of face recognition (Bate & Bennetts, 2014); given that the present study only found face-selective perceptual deficits, interventions aimed at individuals with sub-clinical deficits should also be perceptual in nature. Second, the most successful results are achieved earlier in development, suggesting that the optimum timeframe for the neuronal plasticity of the face recognition system resides in the first few years of life (e.g. Bate & Bennetts, 2014), emphasizing the practical and theoretical importance of detecting atypical performance in childhood. It also appears that face recognition deficits do not develop because of overall poorer skill or inclination, as improving frequency of eye-contact, joint attention, positive experience with others does not seem to improve face recognition abilities in children with DP (Yardley et al., 2008), providing partial support that atypical face recognition deficits may instead be underpinned by the use of atypical strategies in the face processing strategy itself. Longitudinal work and training paradigms are needed to assess these possibilities. At present, the examination of individual eye-movement strategies may offer a more appropriate and reliable way of detecting atypical face recognition skills, provided tasks are designed in a theoretically appropriate way.

4.5. Limitations

Despite making a valuable contribution to the literature, this project has some methodological limitations. First, due to time restrictions, the behavioural data (Experiment 1) was conducted on a relatively small sample size (N = 94) and is slightly underpowered. As this research is on-going, we have conducted a power analysis to illustrate the ideal sample size that can be achieved as we continue to collect participants. Power analyses based on the largest effect size estimate of d = 1.02 from Perez-Roche et al.'s (2017) comparison of small for gestational age and appropriate for gestational age children indicated that a total of 44 participants would be required, split between the appropriate (>10th percentile) and low (<10th percentile) weight for gestational age conditions, to achieve 90% power (two-tailed). However, this estimate relates to only one of our dependent measures (i.e. face perception); concerns have been raised over whether this figure is largely inflated; and our results were correlational in nature. As such, our power calculations were based on 90% power to detect a medium-sized effect (d = 0.15, for F tests), with a total of three predictors. This power analysis indicates that 115 participants are required in total. This sample size affords more than adequate power to detect a medium-sized effect (β = .9) and is over twice as large than the sample size reported in Perez-Roche et al.'s study.

We express greater caution, however, for interpreting the results of the eye-tracking data (Experiment 2) because the final analysis was only carried out on a sample size of 32. It is important to note that out of the original number of participants that took part (N = 51), two participants were excluded for reporting

problems with their vision, and a further seven participants were excluded based on outliers using the 2SD criterion (see *section 3.2* for our justification). However, there is considerable debate within the literature about how and when it is appropriate to remove outliers. The 2SD criterion used here can be considered problematic. Predominantly this is because both the mean and the standard deviation themselves are sensitive to outliers (Leys et al., 2013). Equally, we cannot be sure that extreme dwell times represent participant error or that they represent the actual behaviour of the participants. In contrast, due to the squares basis that linear regressions are calculated upon, outliers could have an over-weighted effect on our results. Critically, the removal of this data did affect the interpretation of our results in some cases, although not *substantively* (see *Appendix B*). Although we believe this action was justified, we also need many more participants to reduce potential noise within the data. A non-linear transformation on the data, with an increased number of participants, could be an alternative option to outlier removal in future analyses. Likewise, the "social scenes" paradigm we used to provide an index of the allocation of visual attention is comprised of 25 pictures (five of which are fillers) and lasts approximately five minutes. Increasing the number of scenes to view is an additional way to reduce noise within participants; this could be tested and applied to future research. Again, it is particularly unusual we didn't observe the typical finding for increased attention to the eyes (e.g. Peterson & Eckstein, 2012), given that we tested individuals from the typical population. Further, as there are individual differences in preferred point of fixation across trials and eye-movement recordings vary between participants (Peterson & Eckstein,

2013), our findings may be better explained by unsystematic noise in the data. Thus, in order to investigate whether adverse perinatal influences affect the face-processing strategy itself, and to comment on whether the examination of individual eye-movement strategies provide a more appropriate and reliable way of detecting atypical face recognition deficits in adults (and children), this project will continue to recruit participants, in relation to the results of our power analysis (detailed above).

Second, the object tasks we administered only tested one object category (i.e. cars). There are several reasons for why cars are used as a common comparison for faces: they are real-world objects; they are three-dimensional; and they have the same first-order structure (Diamond & Carey, 1986), with parts (i.e. body, wheels, doors) all in a fixed relationship to each other (Dennetts et al., 2012). Despite these similarities, one way in which faces and cars are not matched is pre-experimental familiarity. This is somewhat reflected in the finding that the average score for males is often much higher than the average score for females, although this gender bias is not necessarily attributable to the stereotypical male advantage in car expertise (Dennetts et al., 2012) and consistent findings have been observed with bicycles in studies with children (Bennetts et al., 2017). However, it is possible perception and memory related processes impact object categories differently, raising the possibility that not only are there separate recognition systems for faces and objects, but also for different *types* of objects (Towler & Tree, 2018). Given that controlling for pre-experimental familiarity is difficult, future work may benefit from including multiple object tests and also formally investigate the role of pre-experimental

interest and knowledge of the object classes used within such tests.

Third, although the Cambridge Face Memory Test and the Cambridge Face Perception Test are validated measures of unfamiliar face recognition ability, the versions used in this study are comprised only of Caucasian faces. In this study we did not report the ethnicity of participants (as this was not central to our aims or hypotheses), however, this information might have implications for our findings. One own-group bias in face recognition is the own-race bias (ORB), whereby individuals are generally better at recognizing faces of their own race than those of another (e.g. Sporer, 2001). A meta-analysis of 91 independent samples that consisted of approximately 5,000 participants concluded that individuals are 2.23 times more likely to recognize an own-race face than another-race face in the context of a face recognition experiment (Meissner & Brigham, 2001). Similarly, knowledge of cultural effects on perceptual processes, such as initial eye-movements to faces, is limited. Research has shown that the processing of own-race faces is characterized by more active scanning, with a larger number of short fixations, than to other-race faces (Wu, Laeng, Magnussen, 2012). Further, there is diversity in the regions used to extract information from faces across cultures. Specifically, adults from Western cultures tend to focus on the eyes and the mouth region, whereas adults from Eastern cultures tend to fixate on the nose region (Kelly, Miellet, & Caldara, 2010). More recent research suggests that Eastern Asian and Western Caucasian culture groups share similar *initial* eye-movement strategies. When facial stimuli are presented for 1500*ms* or less, both cultural groups fixate on the same featureless region between the eyes and the nose (Or, Peterson, &

Eckstein, 2015). This raises the possibility that (any) differences between cultures might only become apparent after extended or increased exposure to a stimulus.

Presently, if some of our participants were non-Caucasian, they would have been more likely to make errors on the face recognition tasks in experiment 1, and possibly make different fixations in experiment 2 across stimulus presentation, which would have increased noise in the data. It would be informative to record this information so that at the analysis stage, we can see whether there are significant differences between Caucasian and non-Caucasian participants in all tasks. Further, future analyses should investigate how the pattern of fixations changes across stimulus presentation and see whether these changes are related to participants' ethnicity. If this is the case, non-Caucasian participants could be removed, subjected to a separate analysis, or complete other-ethnicity face recognition tests using the Cambridge Face Memory Test format (McKone et al., 2012) at the outset of the study. Although, presently, we cannot specify this information or conduct further analyses.

Finally, our administration of online tests (Experiment 1) may be considered a limitation, although we defend this decision on two accounts. Given that low birth weight for gestational age is a relatively rare population characteristic (7.0% in England; Office for National Statistics, 2016), Web sampling made it more achievable to recruit a large sample. We also believe that, in this case, a university setting (alone) was not the best place to target recruitment, as research consistently reports poorer educational outcomes in low birth weight children when compared to their typical counterparts (e.g. Lahat et

al., 2015; Hack, 2006; Short et al., 2003; Saigal, 2003). Web and lab-based samples have also been shown to yield comparable results for many questionnaires (e.g. Buchanan, 2007), and for performance- based cognitive and perceptual measures in adults (Germine et al., 2012), despite a variety of personal (i.e. lack of diligence) and situational (i.e. quiet space) variables differing between the two (Kraut et al., 2004). Out of all four tests we administered, only the CFMT+ was directly tested in Germine et al.'s (2012) study, raising the possibility that comparable effects may not be observed with the other tests. However, given the matched format and similarity of the tests (in terms of factors such as visual complexity), we believe this is unlikely. Although we provided participants the study links via email, we cannot know for certain that they were completed in the order in which they were presented. Previous studies using similar methods have presented tasks in the same order for all participants (e.g. Bate et al., 2014), and in a counterbalanced order (e.g. Bennetts et al., 2017). Presently, all participants were instructed to complete the tests in the same order; memory tests were completed before perception tests to ensure that exposure to faces used across both tasks did not interfere with the memory responses. A post-hoc visual inspection of the data revealed that the average percentage score for each test reduced after each test presentation (see *Table 16*). This trend could represent actual participant responding, however, as the order of task presentation remained constant for all participants (i.e. CFMT, CCMT, CFPT, Car Perception Test), this trend could also indicate participant fatigue effects or (potential) misunderstanding. The reduction in scores across each test is more consistent with fatigue effects, rather than practice effects or

misunderstanding. The fact that participants scored highest on the first task suggests a high level of understanding, and the consistent reduction most likely represents fatigue. The question of order-effects might be particularly problematic for our current findings, as we report significant effects only for perceptual processes. It is possible this finding is somewhat confounded with the fact perceptual tasks were only completed after memory tasks. Improving the explanatory value of this study by ruling out competing hypotheses would involve counterbalancing the tasks across participants. In this case, counterbalancing would result in 24 combinations (e.g. 1-2-3-4; 1-2-4-3; 1-3-2-4 etc.). If results show no significant differences between the order of presentation conditions, we can place more confidence in our findings.

Table 16. Table to illustrate the task presentation order for each participant and the average % score for each of these tasks.

Task presentation order	**Task name**	**% Score**
1	Cambridge Face Memory Test	78.33
2	Cambridge Car Memory Test	63.54
3	Cambridge Face Perception Test	50.94
4	Car Perception Test	45.71

What remains unclear is whether comparable effects across the Web and the lab will also be observed in children. Although tasks are age-appropriate, issues with task difficulty, understanding, and sustained concentration are likely

to be heightened in child participants. A complete dataset will begin to provide answers to this question; it may be informative to administer tests to children in the lab and on the Web, to directly compare their data in terms of mean performance, performance variance, and internal reliability.

5. Conclusions

In sum, this study found adverse perinatal effects on face perception ability were present in a relatively large adult sample. Our findings are in line with Perez-Roche et al.'s (2017) previous study that found selective impairments of face perception skills remain stable throughout childhood; our results extend this work and suggest adverse perinatal effects on face perception ability do persist into adulthood. Together, these results suggest that, at a sub-clinical level, individuals who experienced a low birth weight for gestational age are at increased risk of face-selective perception deficits throughout development. We further suggest that (by adulthood at least) these atypical face recognition skills are face-selective but it is less clear whether atypical skills are underpinned by atypical processing strategies in the face-processing strategy itself. It also remains unclear whether atypical perception skills in childhood are face-selective, when in development they become detectable and/or face-selective, and whether some individuals "outgrow" difficulties prior to, or during, adulthood. We also highlight the utility of combining behavioural and implicit measures, particularly when testing children and individuals with sub-clinical deficits. Critically, the present finding provides evidence that (at least in some cases) face-selective perceptual deficits remain consistent and persist into adulthood. As such, individuals who have experienced a low birth weight for

gestational age may be suitable candidates for early interventions that address the development of face perception processing abilities.

6. References

Aarnoudse-Moens, C.S., Weisglas-Kuperus, N., van Goudoever, J.B., Oosterlaan, J. (2009). Meta-analysis of neurobehavioral outcomes in very preterm and/or very low birth weight children. Pediatrics, 124(2), 717–728.

Allen, M. C. (2008). Neurodevelopmental outcomes of preterm infants. Current opinion in neurology, 21(2), 123-128.

Anastasi, J. S., & Rhodes, M. G. (2005). An own-age bias in face recognition for children and older adults. Psychonomic Bulletin and Review, 12(6), 1043–1047.

Anderson, P.J., Doyle, L.W. (2003). Neurobehavioral outcomes of school-age children born extremely low birth weight or very preterm in the 1990s. JAMA, 289(24), 3264 –3272.

Andescavage, N. N., du Plessis, A., McCarter, R., Serag, A., Evangelou, I., Vezina, G., & Limperopoulos, C. (2016). Complex trajectories of brain development in the healthy human fetus. Cerebral Cortex, 27(11), 5274-5283.

Auyeung, B., Wheelwright, S., Allison, C., Atkinson, M., Samarawickrema, N., & Baron-Cohen, S. (2007). The children's empathy quotient and systemizing quotient: Sex differences in typical development and in autism spectrum conditions. Journal of Autism and Developmental Disorders, 39(11), 1509–1521.

Ball, G., Boardman, J. P., Rueckert, D., Aljabar, P., Arichi, T., Merchant, N., & Counsell, S. J. (2011). The effect of preterm birth on thalamic and cortical development. Cerebral Cortex, 22(5), 1016-1024.

Ball, G., Boardman, J. P., Rueckert, D., Aljabar, P., Arichi, T.,

Merchant, N., .& Counsell, S. J. (2011). The effect of preterm birth on thalamic and cortical development. Cerebral Cortex, 22(5), 1016-1024.

Baron-Cohen, S., & Wheelwright, S. (2004). The empathy quotient: an investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. Journal of autism and developmental disorders, 34(2), 163-175.

Barton, J., Cherkasova, M., Press, D., Intriligator, J., & O'Connor, M. (2004). Perceptual function in prosopagnosia. Perception, 33, 939–956.

Bate, S. & Tree, J. (2017). The definition and diagnosis of developmental prosopagnosia. The Quarterly Journal of Experimental Psychology. 70, 193–200.

Bate, S. (2013). Face recognition and its disorders. London: Palgrave Macmillan. Bate, S., Bennetts, R., Mole, J. A., Ainge, J. A., Gregory, N. J., Bobak, A. K., & Bussunt, A. (2015). Rehabilitation of face-processing skills in an adolescent with prosopagnosia: Evaluation of an online perceptual training programme. Neuropsychological Rehabilitation, 25(5), 733-762.

Bate, S., & Bennetts, R. (2014). The rehabilitation of face recognition impairments: A critical review and future directions. Frontiers in Human Neuroscience, 8, 491.

Bate, S., Parris. B., Haslam, C., Kay, J. (2010). Socio-emotional functioning and face recognition ability in the normal population. Personality and Individual Differences, 48, 239– 242.

Behrmann, M., Avidan, G., Marotta, J. J., & Kimchi, R. (2005). Detailed exploration of face-related processing in congenital prosopagnosia: 1.

Behavioral findings. Journal of cognitive neuroscience, 17(7), 1130-1149.

Bennetts, R. J., Murray, E., Boyce, T., & Bate, S. (2017). Prevalence of face recognition deficits in middle childhood. The Quarterly Journal of Experimental Psychology, 70(2), 234– 258.

Bhutta, A. T., Cleves, M. A., Casey, P. H., Cradock, M. M., & Anand, K. J. S. (2002). Cognitive and behavioral outcomes of school-aged children who were born preterm: a meta-analysis. Jama, 288(6), 728-737.

Birmingham, E., Bischof, W. F., & Kingstone, A. (2007). Social attention and real- world scenes: The roles of action, competition and social content. Quarterly Journal of Experimental Psychology, 61, 986–998.

Bobak, A. K., Parris, B. A., Gregory, N. J., Bennetts, R. J., & Bate, S. (2017). Eye-movement strategies in developmental prosopagnosia and "super" face recognition. The Quarterly Journal of Experimental Psychology, 70(2), 201-217.

Bowles, D. C., McKone, E., Dawel, A., Duchaine, B., Palermo, R., Schmalzl, L., et al. (2009). Diagnosing prosopagnosia: Effects of ageing, sex, and participant-stimulus ethnic match on the Cambridge Face Memory Test and Cambridge Face Perception Test. Cognitive Neuropsychology, 26, 423–455.

Briscoe, J., Gathercole, S. E., & Marlow, N. (2001). Everyday memory and cognitive ability in children born very prematurely. The Journal of Child Psychology and Psychiatry and Allied Disciplines, 42(6), 749-754.

Bruce, V., & Young, A. (1986). Understanding face recognition. British journal of psychology. 77(3), 305-327.

Brunsdon, R., Coltheart, M., Nickels, L., & Joy, P. (2006).

Developmental prosopagnosia: A case analysis and treatment study. Cognitive Neuropsychology. 23, 822- 840.

Buchanan, T. (2007). Personality testing on the Internet. Oxford handbook of Internet psychology, 447.

Burns, E. J., Bennetts, R. J., Bate, S., Wright, V. C., Weidemann, C. T., & Tree, J. J. (2017). Intact word processing in developmental prosopagnosia. Scientific Reports, 7(1), 210.

Burt, D. M., & Perrett, D. I. (1997). Perceptual asymmetries in judgements of facial attractiveness, age, gender, speech and expression. Neuropsychologia, 35(5), 685-693.

Busigny, T., Joubert, S., Felician, O., Ceccaldi, M., & Rossion, B. (2010). Holistic perception of the individual face is specific and necessary: Evidence from an extensive case study of acquired prosopagnosia. Neuropsychologia, 48(14), 4057–4092.

Bussunt, A. (2014). Rehabilitation of face-processing skills in an adolescent with prosopagnosia: Evaluation of an online perceptual training programme. Neuropsychological Rehabilitation, 1-30.

Campbell, A., & Tanaka, J. W. (2018). Decoupling category level and perceptual similarity in congenital prosopagnosia. Cognitive Neuropsychology, 35(1-2), 63-65.

Carbon, C. C. (2008). Famous faces as icons. The illusion of being an expert in the recognition of famous faces. Perception, 37(5), 801-806.

Carey, S. and Diamond, R. (1994) Are faces perceived as configurations more by adults than by children? Visual Cognition, 1, 253–274.

Carey, S., & Diamond, R. (1977). From piecemeal to configurational representation of faces. Science. 195, 312-314.

Chen, J. H., Claessens, A., & Msall, M. E. (2014). Prematurity and school readiness in a nationally representative sample of Australian children: does typically occurring preschool moderate the relationship?. Early human development, 90(2), 73-79.

Collishaw, S. M., & Hole, G. J. (2000). Featural and configurational processes in the recognition of faces of different familiarity. Perception, 29(8), 893-909.

Colombo, J., & Mitchell, D. W. (1990). Individual differences in early visual attention: Fixation time and information processing. Individual differences in infancy: Reliability, stability, prediction, 193-227.

Colombo, J., Mitchell, D. W., Coldren, J. T., & Freeseman, L. J. (1991). Individual differences in infant visual attention: Are short lookers faster processors or feature processors?. Child development, 62(6), 1247-1257.

Crawford, J. R., Garthwaite, P. H., & Gray, C. D. (2003). Wanted: Fully operational definitions of dissociations in single-case studies. Cortex. 39, 357–370.

Crookes, L., & McKone, E. (2009). Early maturity of face recognition: No childhood development of holistic processing, novel face encoding, or face-space. Cognition, 11, 219– 247.

Croydon, A., Pimperton, H., Ewing, L., Duchaine, BC., Pellicano, E. (2014). The Cambridge Face Memory Test for Children (CFMT-C): a new tool for measuring face recognition skills in childhood. Neuropsychologia, 62, 60–

67.

Dalrymple, K. A., Elison, J. T., & Duchaine, B. (2017). Face-specific and domain- general visual processing deficits in children with developmental prosopagnosia. The Quarterly Journal of Experimental Psychology, 1-17.

Dalrymple, K. A., Gomez, J., & Duchaine, B. (2012). CFMT-Kids: A new test of face memory for children. Journal of Vision. 12(9), 492-492.

Dalrymple, K.A., Garrido, L., Duchaine, B. (2014). Dissociation between face perception and face memory in adults, but not children, with developmental prosopagnosia. Developmental Cognitive Neuroscience. 10, 10–20.

de Heering, A., Rossion, B., Maurer, D. (2012). Developmental changes in face recognition during childhood: evidence from upright and inverted faces. Cognitive Development, 27, 17–27.

DeGutis, J., Cohan, S., & Nakayama, K. (2014). Holistic face training enhances face processing in developmental prosopagnosia. Brain, 137(6), 1781–1798.

DeGutis, J.M., Bentin, S., Robertson, L.C., & D'Esposito, M. (2007). Functional plasticity in ventral temporal cortex following cognitive rehabilitation of a congenital prosopagnosia. Journal of Cognitive Neuroscience, 19, 1790–1802.

Dennett, H. W., McKone, E., Tavashmi, R., Hall, A., Pidcock, M., Edwards, M., & Duchaine, B. (2012). The Cambridge Car Memory Test: A task matched in format to the Cambridge Face Memory Test, with norms, reliability, sex differences, dissociations from face memory, and expertise effects.

Behavior Research Methods, 44(2), 587-605.

Diamond, R., & Carey, S. (1986). Why faces are and are not special: An effect of expertise. Journal of Experimental Psychology: General, 115, 107-1 17.

Duchaine, B., & Yovel, G. (2015). A revised neural framework for face processing. Annual Review of Vision Science, 1, 393-416.

Duchaine, B., Nakayama, K. (2006b). The Cambridge Face Memory Test: results for neurologically intact individuals and an investigation of its validity using inverted face stimuli and prosopagnosic participants. Neuropsychologia, 44, 576–585.

Duchaine, B.C., Nakayama, K. (2004). Developmental prosopagnosia and the Benton facial recognition test. Neurology. 62, 1219–1220.

Eimer, M. (2018). What do associations and dissociations between face and object recognition abilities tell us about the domain-generality of face processing? Cognitive Neuropsychology. 35(1-2), 80-82.

Elgen, I., Sommerfelt, K., Markestad, T. (2002) Population based, controlled study of behavioural problems and psychiatric disorders in low birthweight children at 11 years of age. Archives of Disease in Childhood Fetal and Neonatal Edition. 87, 128-132.

Fenoglio, A., Georgieff, M. K., & Elison, J. T. (2017). Social brain circuitry and social cognition in infants born preterm. Journal of neurodevelopmental disorders, 9(1), 27.

Figueras, F., Eixarch, E., Gratacos, E., & Gardosi, J. (2008). Predictiveness of antenatal umbilical artery Doppler for adverse pregnancy

outcome in small-for-gestational-age babies according to customised birthweight centiles: population-based study. BJOG: An International Journal of Obstetrics & Gynaecology, 115(5), 590-594.

Firestone, A., Turk-Browne, N. B., & Ryan, J. D. (2007). Age-related deficits in face recognition are related to underlying changes in scanning behavior. Aging, Neuropsychology, and Cognition, 14(6), 594-607.

Freire, A., & Lee, K. (2001). Face recognition in 4- to 7-year-olds: Processing of configural, featural, and paraphernalia information. Journal of Experimental Child Psychology. 80, 347 – 371.

Frie, J., Padilla, N., Ådén, U., Lagercrantz, H., & Bartocci, M. (2016). Extremely preterm-born infants demonstrate different facial recognition processes at 6-10 months of corrected age. The Journal of pediatrics, 172, 96-102.

Fumagalli, M., Provenzi, L., De Carli, P., Dessimone, F., Sirgiovanni, I., Giorda, R., & Brambilla, P. (2018). From early stress to 12-month development in very preterm infants: Preliminary findings on epigenetic mechanisms and brain growth. PloS one, 13(1).

Garrido, L., Duchaine, B., & DeGutis, J. (2018). Association vs dissociation and setting appropriate criteria for object agnosia. Cognitive Neuropsychology. 35(1–2), 55–58.

Gauthier, I., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. Nature neuroscience, 3(2), 191.

Geldart, S., Mondloch, C. J., Maurer, D., De Schonen, S., & Brent, H. P.

(2002). The effect of early visual deprivation on the development of face processing. Developmental Science, 5(4), 490-501.

Geldart, S., Mondloch, C., Maurer, D., de Schonen, S., & Brent, H. (2002). The effect of early visual deprivation on the development of face processing. Developmental Science, 5, 490–501.

Germine, L., Cashdollar, N., Düzel, E., & Duchaine, B. (2011a). A new selective developmental deficit: Impaired object recognition with normal face recognition. Cortex, 47, 598–607.

Germine, L., Nakayama, K., Duchaine, B. C., Chabris, C. F., Chatterjee, G., & Wilmer, J. B. (2012). Is the web as good as the lab? Comparable performance from Web and lab in cognitive/perceptual experiments. Psychonomic Bulletin and Review, 19, 847–857.

Geskin, J., & Behrmann, M. (2018). Congenital prosopagnosia without object agnosia? A literature review. Cognitive Neuropsychology. 35(1–2), 4–54.

Goldberger, A. S., & Goldberger, A. S. G. (1991). *A course in econometrics*. Harvard University Press.

Gomez, J., Pestilli, F., Witthoft, N., Golarai, G., Liberman, A., Poltoratski, S., & Grill-Spector, K. (2015). Functionally defined white matter reveals segregated pathways in human ventral temporal cortex associated with category-specific processing. Neuron, 85(1), 216-227.

Gosselin, F., & Schyns, P. G. (2001). Bubbles: A technique to reveal the use of information in recognition tasks. Vision Research, 41(17), 2261–2271.

Gray, C. M., & Cook, R. (2018). Should developmental prosopagnosia, developmental body agnosia, and developmental object agnosia be considered

independent neurodevelopmental conditions? Cognitive Neuropsychology, 35 (1-2), 59-62.

Hack, M. (2006) Young adult outcomes of very-low-birth weight children. Seminars in Fetal and neonatal Medicine, 11, 127–37.

Hack, M., Flannery, DJ., Schluchter, M., Cartar, L., Borawski, E., & Klein, N. (2002). Outcomes in Young Adulthood for Very-Low-Birth weight Infants. New England Journal of Medicine, 246(3), 149-157.

Hall, J. K., Hutton, S. B., & Morgan, M. J. (2010). Sex differences in scanning faces: Does attention to the eyes explain female superiority in facial expression recognition?. Cognition & Emotion, 24(4), 629-637.

Hills, P. J., & Lewis, M. B. (2011). The own-age face recognition bias in children and adults. The Quarterly Journal of Experimental Psychology, 64, 17–23.

Hills, P. J., & Willis, S. F. (2016). Children view own-age faces qualitatively differently to other-age faces. Journal of Cognitive Psychology, 28(5), 601-610.

Hole, G. J. (1994). Configurational factors in the perception of unfamiliar faces. Perception, 23(1), 65-74.

Hsiao, J. H. W., & Cottrell, G. (2008). Two fixations suffice in face recognition. Psychological Science, 19(10), 998-1006.

Indredavik, M.S., Vik, T., Heyerdahl, S., Kulseng, S., Fayers, P., Brubakk, A.M. (2004). Psychiatric symptoms and disorders in adolescents with low birth weight. Arch Dis Child Fetal Neonatal Ed. 89, 445–450.

Innocenti, G. M., & Price, D. J. (2005). Exuberance in the development

of cortical networks. Nature Reviews Neuroscience, 6(12), 955.

Johnson, M. H. (2011). Interactive specialization: a domain-general framework for human functional brain development?. Developmental cognitive neuroscience, 1(1), 7-21.

Jones, R. D., & Tranel, D. (2001). Severe developmental prosopagnosia in a child with superior intellect. Journal of Clinical and Experimental Neuropsychology. 23, 265–273.

Kanwisher, N., McDermott, J., & Chun, M.M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *Journal of neuroscience*. 17(11), 4302-4311.

Karmiloff-Smith, B. A. (1994). Beyond modularity: A developmental perspective on cognitive science. European journal of disorders of communication, 29(1), 95-105.

Kelly, D. J., Miellet, S., & Caldara, R. (2010). Culture shapes eye movements for visually homogeneous objects. Frontiers in psychology, 1, 6.

Kennerknecht, I., Grueter, T., Welling, B., Wentzek, S., Horst, J., Edwards, S., & Grueter, M. (2006). First report of prevalence of non-syndromic hereditary prosopagnosia (HPA). American Journal of Medical Genetics Part A. 140(15), 1617-1622.

Kracke, I. (1994). Developmental prosopagnosia in Asperger syndrome: presentation and discussion of an individual case. Developmental Medicine and Child Neurology. 36, 873–886.

Kraha, A., Turner, H., Nimon, K., Zientek, L., & Henson, R. (2012). Tools to support interpreting multiple regression in the face of

multicollinearity. Frontiers in psychology, 3, 44.

Kraut, R., Olson, J., Banaji, M., Bruckman, A., Cohen, J., & Couper, M. (2004). Psychological research online: report of Board of Scientific Affairs' Advisory Group on the Conduct of Research on the Internet. American psychologist, 59(2), 105.

Lahat, A., Van Lieshout, R. J., Saigal, S., Boyle, M. H., & Schmidt, L. A. (2014). ADHD among young adults born at extremely low birth weight: The role of fluid intelligence in childhood. Frontiers in Psychology, 5, 446.

Le Grand, R., Mondloch, C. J., Maurer, D. & Brent, H. P. (2003). Expert face processing requires visual input to the right hemisphere during infancy. Nature Neuroscience, 6, 1108–12.

Le Grand, R., Mondloch, C. J., Maurer, D., & Brent, H. P. (2004). Impairment in holistic face processing following early visual deprivation. Psychological Science, 15, 762– 768.

Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. Journal of Experimental Social Psychology, 49(4), 764-766.

Li, J., Tian, M., Fang, H., Xu, M., Li, H., & Liu, J. (2010). Extraversion predicts individual differences in face recognition. Communicative & Integrative Biology, 3, 295-298.

Linsell, L., Malouf, R., Morris, J., Kurinczuk, J.J., Marlow, N. (2017). Risk factor models for Neurodevelopmental outcomes in children born very preterm or with very low birth weight: a systematic review of methodology and reporting. American Journal of Epidemiology, 185, 601–12.

Liversedge, S. P., & Findlay, J. M. (2000). Saccadic eye movements and cognition. Trends in cognitive sciences, 4(1), 6-14.

Lundequist, A., Bohm, B., Lagercrantz, H., Forssberg, H., Smedler, A.C. (2014). Cognitive outcome varies in adolescents born preterm, depending on gestational age, intrauterine growth and neonatal complications. Acta Paediatr, 104(3), 292–9.

Luu, T. M., Vohr, B. R., Allan, W., Schneider, K. C., & Ment, L. R. (2011). Evidence for catch-up in cognition and receptive vocabulary among adolescents born very preterm. Pediatrics, 2010.

Luu, T.M., Ment, L., Allan, W., Schneider, K., & Vohr, B.R. (2011). Executive and memory function in adolescents born very preterm. Pediatrics, 127(3), 639–646.

Ma, Y., & Han, S. (2010). Why we respond faster to the self than to others? An implicit positive association theory of self-advantage during implicit face recognition. Journal of Experimental Psychology: Human Perception and Performance, 36(3), 619.

McKone, E., Crookes, K., Jeffery, L., & Dilks, D. D. (2012). A critical review of the development of face recognition: Experience is less important than previously believed. Cognitive neuropsychology, 29(1-2), 174-212.

McKone, E., Stokes, S., Liu, J., Cohan, S., Fiorentini, C., Pidcock, M., & Pelleg, M. (2012). A robust method of measuring other-race and other-ethnicity effects: The Cambridge Face Memory Test format. *PLoS One, 7*(10), e47956.

Megreya, A. M., & Burton, A. M. (2006). Unfamiliar faces are not

faces: Evidence from a matching task. Memory & cognition, 34(4), 865-876.

Meissner, C. A., & Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. Psychology, Public Policy, and Law, 7(1), 3.

Mento, G., & Bisiacchi, P. S. (2012). Neurocognitive development in preterm infants: insights from different approaches. Neuroscience & Biobehavioral Reviews, 36(1), 536-555.

Mograss, MA., Guillem, F., Stickgold, R. (2010). Individual differences in face recognition memory: comparison among habitual short, average, and long sleepers. Behavioural Brain Research, 208, 576–583.

Natu, V., & O'Toole, A. J. (2011). The neural processing of familiar and unfamiliar faces: a review and synopsis. British Journal of Psychology, 102(4), 726-747.

Nelson CA. (2001). The development and neural bases of face recognition. Infant and Child Development, 10, 3–18.

O'Connor, A. R., Stephenson, T. J., Johnson, A., Tobin, M. J., Ratib, S., Moseley, M., & Fielder, A. R. (2004). Visual function in low birthweight children. British Journal of Ophthalmology, 88(9), 1149-1153.

Office for National Statistics (2016). Birth Characteristics in England and Wales: /bulletins/birthcharacteristicsinenglandandwales/2016 (accessed 7th February 2018).

Or, C. C. F., Peterson, M. F., & Eckstein, M. P. (2015). Initial eye movements during face identification are optimal and similar across cultures. Journal of vision, 15(13), 12-12.

Palermo, R., Willis, M. L., Rivolta, D., McKone, E., Wilson, C. E., & Calder, A. J. (2011). Impaired holistic coding of facial expression and facial identity in congenital prosopagnosia. Neuropsychologia, 49(5), 1226-1235.

Perez-Roche, T., Altemir, I., Giménez, G., Prieto, E., Gonzalez, I., Pisón, J. L., & Pueyo, V. (2017). Face recognition impairment in small for gestational age and preterm children. Research in developmental disabilities, 62, 166-173.

Perlman, S. B., Morris, J. P., Vander Wyk, B. C., Green, S. R., Doyle, J. L., and Pelphrey, K. A. (2009). Individual differences in personality predict how people look at faces. PLoS ONE, 4, 5952.

Peterson, B. S., Vohr, B., Staib, L. H., Cannistraci, C. J., Dolberg, A., Schneider, K. C., & Duncan, C. C. (2000). Regional brain volume abnormalities and long-term cognitive outcome in preterm infants. JAMA, 284(15), 1939-1947.

Peterson, M. F., & Eckstein, M. P. (2012). Looking just below the eyes is optimal across face recognition tasks. Proceedings of the National Academy of Sciences, 109(48), 3314-3323.

Peterson, M.F., Eckstein, M.P. (2013) Individual differences in eye movements during face identification reflect observer-specific optimal points of fixation. Psychological Science, 24, 1216 – 1225.

Pitcher, D., Walsh, V., & Duchaine, B. (2011). The role of the occipital face area in the cortical face perception network. Experimental brain research, 209(4), 481-493.

Pueyo, V., Oros, D., Valle, S., Tuquet, H., Güerri, N., Argüelles, M., &

Ventura, P. (2012). Axonal loss and cognitive deficits in term infants with normal umbilical artery Doppler born small-for-gestational age. Ultrasound in Obstetrics & Gynecology, 40(3), 297-303.

Ramus, F. (2004). Neurobiology of dyslexia: A reinterpretation of the data. Trends in Neurosciences, 27, 720– 726.

Reynolds, C. R., & Voress, J. K. (2007). Test of Memory and Learning (TOMAL–2) (2nd ed.). Austin, TX: Pro-Ed.

Rhodes, G., Jeffery, L., Taylor, L., Hayward, W. G., & Ewing, L. (2014). Individual differences in adaptive coding of face identity are linked to individual differences in face recognition ability. Journal of Experimental Psychology: Human Perception & Performance, 40, 897–903.

Richler, J. J., Cheung, O. S., & Gauthier, I. (2011). Holistic processing predicts face recognition. Psychological Science, 22, 464–471.

Rose, S. A., Feldman, J. F., & Jankowski, J. J. (2001). Attention and recognition memory in the 1st year of life: a longitudinal study of preterm and full-term infants. Developmental psychology, 37(1), 135.

Rose, S. A., Feldman, J. F., Jankowski, J. J., & Van Rossem, R. (2005). Pathways from prematurity and infant abilities to later cognition. Child development, 76(6), 1172-1184.

Rosenthal, G., & Avidan, G. (2018). A possible neuronal account for the behavioral heterogeneity in congenital prosopagnosia. Cognitive Neuropsychology, 35 (1-2), 74-77.

Rossion, B., Dricot, L., Devolder, A., Bodart, J.M., Crommelinck, M., De Gelder, B., & Zoontjes, R. (2000). Hemispheric asymmetries for whole-

based and part-based face processing in the human fusiform gyrus. Journal of Cognitive Neuroscience, 12, 793–802.

Russell, R., Duchaine, B., Nakayama, K. (2009). Super-recognizers: People with extraordinary face recognition ability. Psychonomic Bulletin and Review, 16, 252–257.

Saigal, S., den Ouden, L., Wolke, D., Hoult, L., Paneth, N., Streiner, D. L., & Pinto-Martin, J. (2003). School-age outcomes in children who were extremely low birth weight from four international population-based cohorts. Pediatrics, 112(4), 943-950.

Schiltz, C., & Rossion, B. (2006). Faces are represented holistically in the human occipitotemporal cortex. Neuroimage, 32, 1385–1394.

Schmalzl, L., Palermo, R., Green, M., Brunsdon, R., & Coltheart, M. (2008). Training of familiar face recognition and visual scan paths for faces in a child with congenital prosopagnosia. Cognitive Neuropsychiatry, 25, 704–729.

Schultz, R. T. (2005). Developmental deficits in social perception in autism: The role of the amygdala and fusiform face area. International Journal of Developmental Neuroscience, 23, 125– 141.

Schwarzer, G., Huber, S., Grueter, M., Grueter, T., Grob, C., Hipfel, M., & Kennerknecht, I. (2007). Gaze behaviour in hereditary prosopagnosia. Psychological Research, 71, 583– 590.

Short, E. J., Klein, N. K., Lewis, B. A., Fulton, S., Eisengart, S., Kercsmar, C., & Singer, L. T. (2003). Cognitive and academic consequences of bronchopulmonary dysplasia and very low birth weight: 8-year-old outcomes. Pediatrics, 112(5), 359-359.

Smith, M. L., Gosselin, F., & Schyns, P. G. (2004). Receptive fields for flexible face categorizations. Psychological Science, 15(11), 753-761.

Sporer, S. L. (2001). Recognizing faces of other ethnic groups: An integration of theories. Psychology, Public Policy, and Law, 7(1), 36.

Sripada, K., Løhaugen, G.C., Eikenes, L., Bjørlykke, K.M., Håberg, A.K., Skranes, J., & Rimol, L.M. (2015). Visual-motor deficits relate to altered gray and white matter in young adults born preterm with very low birth weight. NeuroImage, 109, 493–504.

Starrfelt, R., & Robotham, R. J. (2018). On the use of cognitive neuropsychological methods in developmental disorders: A commentary on Geskin and Behrmann. Cognitive Neuropsychology, 35(1-2), 94-97.

Stein, R.E., Siegel, M.J., & Bauman, L.J. (2006) Are children of moderately low birth weight at increased risk for poor health? A new look at an old question. Pediatrics, 118, 217–223.

Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. The Quarterly Journal of Experimental Psychology Section A, 46(2), 225-245.

Tavor, I., Yablonski, M., Mezer, A., Rom, S., Assaf, Y., & Yovel, G. (2014). Separate parts of occipito-temporal white matter fibers are associated with recognition of faces and places. Neuroimage, 86, 123-130.

Taylor, M.J., Edmonds, G.E., McGarthy, G., & Allison, T. (2001). Eyes first! Eye processing develops before face processing in children. NeuroReport, 12, 1671–1676.

Thomas, C., Moya, L., Avidan, G., Humphreys, K., Jung, K. J.,

Peterson, M. A., & Behrmann, M. (2008). Reduction in white matter connectivity, revealed by diffusion tensor imaging, may account for age-related changes in face perception. Journal of Cognitive Neuroscience, 20(2), 268-284.

Thompson, D. K., Warfield, S. K., Carlin, J. B., Pavlovic, M., Wang, H. X., Bear, M., & Inder, T. E. (2006). Perinatal risk factors altering regional brain structure in the preterm infant. Brain, 130(3), 667-677.

Towler, J.R., & Tree, J.J. (2018). Commonly associated face and object recognition impairments have implications for the cognitive architecture. Cognitive Neuropsychology, 35(1-2), 70-73.

Turati, C., Di Giorgio, E., Bardi, L., & Simion, F. (2010). Holistic face processing in newborns, 3-month-old infants, and adults: Evidence from the composite face effect. Child Development, 81(6), 1894–1905.

Turken, U., Whitfield-Gabrieli, S., Bammer, R., Baldo, J. V., Dronkers, N. F., & Gabrieli, J. D. (2008). Cognitive processing speed and the structure of white matter pathways: convergent evidence from normal variation and lesion studies. Neuroimage, 42(2), 1032-1044.

Unger, A., Alm, K. H., Collins, J. A., O'Leary, J. M., & Olson, I. R. (2016). Variation in white matter connectivity predicts the ability to remember faces and discriminate their emotions. Journal of the International Neuropsychological Society, 22(2), 180-190.

United Nations Children's Fund and World Health Organization, Low Birthweight: Country, regional and global estimates. (2004). UNICEF, New York.

(http://www.who.int/reproductivehealth/publications/low_birthweight/low_birth

weight_estim ates.pdf)

Van Baar, A. L., Ultee, K., Gunning, W. B., Soepatmi, S., & de Leeuw, R. (2006). Developmental course of very preterm children in relation to school outcome. Journal of Developmental and Physical Disabilities, 18(3), 273-293.

Vinette, C., Gosselin, F., & Schyns, P. G. (2004). Spatio-temporal dynamics of face recognition in a flash: It's in the eyes. *Cognitive Science*, *28*(2), 289-301.

Wang, R., Li, J., Fang, H., Tian, M., & Liu, J. (2012). Individual Differences in Holisitic Processing Predict Face Recognition Ability. Psychological Science, 23(2), 169 – 177.

Want, S.C., Pascalis, O., Coleman, M., & Blades, M. (2003). Recognizing people from the inner and outer parts of their faces: Developmental data concerning 'unfamiliar' faces. British Journal of Developmental Psychology, 21, 125–135.

Weigelt, S., Koldewyn, K., & Kanwisher, N. (2012). Face identity recognition in autism spectrum disorders: a review of behavioral studies. Neuroscience & Biobehavioral Reviews, 36(3), 1060-1084.

Weigelt, S., Koldewyn, K., Dilks, D.D., Balas, B., McKone, E., Kanwisher, N. (2014). Domain-specific development of face memory but not face perception. Developmental Science, 17, 47–58.

Weiss, N., Mardo, E., & Avidan, G. (2016). Visual expertise for horses in a case of congenital prosopagnosia. Neuropsychologia, 83, 63–75.

Wheelwright, S., Baron-Cohen, S., Goldenfeld, N., Delaney, J., Fine, D., & Smith, R. (2006). Predicting Autism Spectrum Quotient (AQ) from the

Systemizing Quotient- Revised (SQ-R) and Empathy Quotient (EQ). Brain Research, 1079(1), 47–56.

Wilmer, J. B., Germine, L., Chabris, C. F., Chatterjee, G., Williams, M., & Loken, E. (2010). Human face recognition ability is specific and highly heritable. Proceedings of the National Academy of Sciences, 107, 5238–5241.

Wu, E. X. W., Laeng, B., & Magnussen, S. (2012). Through the eyes of the own-race bias: Eye-tracking and pupillometry during face recognition. Social neuroscience, 7(2), 202-216.

Xu, Y. et al. (2005) The M170 is selective for faces, not for expertise. Neuropsychologia, 43, 588–597.

Yardley, L., McDermott, L., Pisarski, S., Duchaine, B., Nakayama, K. (2008). Psychosocial consequences of developmental prosopagnosia: A problem of recognition. Journal of Psychosomatic Research, 65(5), 445– 451.

Yin, R.K. (1969) Looking at upside-down faces. Journal of Experimental Psychology, 81, 141–145.

Yovel, G., & Kanwisher, N. (2004). Face perception: Domain specific, not process specific. Neuron, 44, 889–898.

Zhao, Y., Li, J., Liu, X., Song, Y., Wang, R., Yang, Z., & Liu, J. (2016). Altered spontaneous neural activity in the occipital face area reflects behavioral deficits in developmental prosopagnosia. Neuropsychologia, 89, 344–355.

Zoccolotti, P., De Luca, M., Di Pace, E., Gasperini, F., Judica, A., & Spinelli, D. (2005). Word length effect in early reading and in developmental dyslexia. Brain and language, 93(3), 369-373.

7. Appendices

7.1. Appendix A – **comparing CFMT+ and CFMT results for Experiment 1**

7.1.1. Correlations

Table A. Correlation coefficients of birth weight, gestation, and percentile with the CFMT+ and the CFMT. p-values are reported in parentheses and significant correlations are highlighted in bold. Correlations indicating a trend towards significance are italicised.

	Hypothesised Predictors		
	Birth weight	**Gestation**	**Percentile**
CFMT+	**.207 (.047)**	.136 (.194)	.177 (.089)[a]
CFMT	*.201 (.053)*	.088 (.401)	**.212 (.042)[a]**

7.1.2. Multiple linear regressions

Birthweight

Table B. Summary of the multiple linear regression analysis predicting birthweight from CFMT, CCMT, CFPT, CCPT, EQ and SQ scores.

Variable	***t***	**b**	**β**	***F***	R^2	Adj. R^2	***p***
Overall Model				1.076	.070	.005	.383
CFMT Scores	1.610	.011	.184				.111
CCMT Scores	-.322	-.002	-.038				.748
CFPT Scores	1.342	.011	.151				.183
CCPT Scores	.355	.002	0.38				.723
EQ Scores	-.451	-.003	.012				.653
SQ Scores	.115	.001	-.048				.908

Note. CFMT = Cambridge Face Memory Test, CCMT = Cambridge Car Memory Test, CFPT = Cambridge Face Perception Test, CCPT = Cambridge Car Perception Test, EQ = Empathy Quotient, SQ = Systemizing Quotient.

Gestation period

Table C. Summary of the multiple linear regression analysis predicting gestation period from CFMT, CCMT, CFPT, CCPT, EQ and SQ scores.

Variable	*t*	b	β	*F*	R^2	Adj. R^2	*p*
Overall Model				0.448	.030	-.037	.844
CFMT Scores	1.180	.031	.138				.241
CCMT Scores	.417	.011	.051				.678
CFPT Scores	-.390	-.013	-.045				.697
CCPT Scores	.345	.006	.038				.731
Empathy Quotient	-.511	-.012	-.055				.611
Systemizing Quotient	.504	.010	.054				.616

Note. CFMT = Cambridge Face Memory Test, CCMT = Cambridge Car Memory Test, CFPT = Cambridge Face Perception Test, CCPT = Cambridge Car Perception Test, EQ = Empathy Quotient, SQ = Systemizing Quotient.

Percentile

Table D. Summary of the multiple linear regression analysis predicting percentile from CFMT, CCMT, CFPT, CCPT, EQ and SQ scores.

Variable	*t*	b	β	*F*	R^2	Adj. R^2	*p*
Overall Model				1.813	.112	.050	.106
CFMT Scores	1.126	.320	.126				.263
CCMT Scores	-.479	-.132	-.056				.633
CFPT Scores	2.537	.886	.280				.013*
CCPT Scores	-.788	-.154	-.083				.433
Empathy Quotient	.133	.033	-.075				.894
Systemizing Quotient	-.735	-.161	.014				.465

Note. CFMT = Cambridge Face Memory Test, CCMT = Cambridge Car Memory Test, CFPT = Cambridge Face Perception Test, CCPT = Cambridge Car Perception Test, EQ = Empathy Quotient, SQ = Systemizing Quotient.

** $p < .05$.*

7.2. Appendix B – **Experiment 2 summary without outliers removed**

7.2.1. Normality tests

When outliers were not removed, normality tests remained largely similar. When outliers were removed, percentile, left eye dwell time, and overall eye dwell time were normally distributed; when outliers were not removed, percentile, left eye dwell time and overall eye dwell time deviated significantly from normality.

Table E. Summary table of normality tests for the nine dependent variables in experiment 2 when outliers are not removed. Significant p-values are marked with an asterisk (*).

	Shapiro-Wilk test		
	Statistic	df	*p*-value
Birth weight	.905	42	.002*
Gestation	.794	42	.000*
Percentile	.946	42	.046*
Body Dwell Time	.957	42	.112
Face Dwell Time	.981	42	.687
Hair Dwell Time	.954	42	.092
Inner Features Dwell Time	.975	42	.469
Eyes Dwell Time	.943	42	.038*
Nose Dwell Time	.974	42	.440
Mouth Dwell Time	.925	42	.009*
Left Eye Dwell Time	.938	42	.024*
Right Eye Dwell Time	.893	42	.001*

7.2.2. Correlations

Correlations were re-run without the outliers removed. This produced no significant correlations and no trends towards significance (see *Table F*).

Table F. Correlation coefficients of birth weight, gestation, and percentile with dwell time (%) on bodies, hair, faces, and inner features, including the eyes (overall and individual), nose and mouth, when outliers are not removed. p-values are reported in parentheses and significant correlations are highlighted in bold.

	Hypothesised Predictors		
	Birth weight	Gestation	Percentile
Body Dwell Time	.059 *(.711)*	-.135 *(.369)*	.001 *(.996)*
Face Dwell Time	-.020 *(.901)*	.032 *(.841)*	.069 *(.666)*
Hair Dwell Time	.034 *(.830)*	-.053 *(.739)*	-.054 *(.733)*
Inner Features Dwell Time	-.045 *(.778)*	.040 *(.803)*	.016 *(.919)*
Eyes Dwell Time	-.208 *(.187)*	.114 *(.473)*	-.191 *(.225)*
Nose Dwell Time	-.007 *(.967)*	-.032 *(.840)*	.061 *(.700)*
Mouth Dwell Time	.176 *(.266))*	-.068 *(.671)*	.162 *(.307)*
Left Eye Dwell Time	-.103 *(.518)*	.016 *(.918)*	-.081 *(.612)*
Right Eye Dwell Time	-.233 *(.137)*	.152 *(.337)*	-.231 *(.141)*

7.2.3. Multiple linear regressions

Regressions were re-run without the outliers removed.

Birthweight

The birthweight regression analysis predicting birthweight from body dwell time, hair dwell time, and inner features dwell time without the outliers removed outputted similar results, and associated conclusions. The model and all the predictors were non-significant (see *Table G*). It is important to note, however, that when outliers were removed there were trends towards significance; these trends were not observed when outliers were included in analysis.

Table G. Summary of the multiple linear regression analysis predicting birthweight from body dwell time, hair dwell time, and inner features dwell time, when the outliers are not removed.

Variable	***t***	**b**	**β**	***F***	R^2	**Adj.** R^2	***p***
Overall Model				.140	.011	-.067	.935
Body Dwell Time	.266	3.198	.079				.791
Hair Dwell Time	.355	9.919	.062				.724
Inner Features Dwell Time	-.015	-.131	-.004				.988

The birthweight regression analysis predicting birthweight from left eye dwell time, right eye dwell time, nose dwell time, and mouth dwell time without the outliers removed outputted differing results, and associated conclusions than

when the outliers were removed. The model and all the predictors became non-significant (see *Table H*). Right eye dwell time did show a trend towards significance. This is in-keeping with the results obtained without the outliers removed, and may just reflect a lack of power.

Table H. Summary of the multiple linear regression analysis predicting birthweight from left eye dwell time, right eye dwell time, nose dwell time, and mouth dwell time, when the outliers are not removed.

Variable	***t***	**b**	**β**	***F***	R^2	**Adj.** R^2	***p***
Overall Model				1.431	.134	.040	.243
Left Eye Dwell Time	.733	15.722	.160				.468
Right Eye Dwell Time	-1.916	-26.269	-.430				.063
Nose Dwell Time	-.042	-.309	-.007				.966
Mouth Dwell Time	.142	1.234	.029				.888

Gestation Period

The gestation period regression analysis predicting gestation from body dwell time, hair dwell time, and inner features dwell time without the outliers removed outputted similar results, and associated conclusions. The model and all the predictors were non-significant (see *Table I*).

Table I. Summary of the multiple linear regression analysis predicting gestation period from body dwell time, hair dwell time, and inner features dwell time, when the outliers are not removed.

Variable	*t*	b	β	*F*	R^2	Adj. R^2	*p*
Overall Model				.461	.035	-.041	.711
Body Dwell Time	-.298	-17.028	-.087				-.298
Hair Dwell Time	.982	130.633	.170				.982
Inner Features Dwell Time	-.098	-4.166	-.029				-.098

The gestation period regression analysis predicting gestation from left eye dwell time, right eye dwell time, nose dwell time, and mouth dwell time without the outliers removed produced comparable results, and associated conclusions (see *Table J*).

Table J. Summary of the multiple linear regression analysis predicting gestation period from left eye dwell time, right eye dwell time, nose dwell time, and mouth dwell time, when outliers are not removed.

Variable	*t*	b	β	*F*	R^2	Adj. R^2	*p*
Overall Model				.179	.019	-.087	.948
Left Eye Dwell Time	.178	84.146	.178				.450
Right Eye Dwell Time	-.075	-22.259	-.075				.754
Nose Dwell Time	-.067	-14.115	-.067				.709
Mouth Dwell Time	-.015	-2.976	-.015				.947

Percentile

The percentile regression analysis predicting percentile from body dwell time, hair dwell time, and inner features dwell time without the outliers removed outputted differing results, and associated conclusions. The model and all the predictors became non-significant (see *Table K*). Specifically, inner features dwell time was significant with outliers removed but was non-significant when outliers were included in the analysis. A trend towards significance was observed in the case of body dwell time when outliers were removed but this was relationship was abolished when outliers were included in the analysis.

Table K. Summary of the multiple linear regression analysis predicting percentile from body dwell time, hair dwell time, and inner features dwell time when outliers are not removed.

Variable	***t***	**b**	**β**	***F***	R^2	**Adj.** R^2	***p***
Overall Model				.094	.007	-.071	.963
Body Dwell Time	-.279	-149.310	-.082				.782
Hair Dwell Time	-.474	-590.122	-.083				.638
Inner Features Dwell Time	-.181	-72.116	-.055				.857

The percentile regression analysis predicting percentile from left eye dwell time, right eye dwell time, nose dwell time, and mouth dwell time without the outliers removed produced differing results, and associated conclusions. The model and all the predictors became non-significant (see *Table L*). Specifically,

right eye dwell time was significant with the outliers removed but was non-significant when outliers were included in the analysis. A trend towards significance was observed in the case of left eye dwell time when outliers were removed but this effect was abolished when outliers were included in the analysis.

Table L. Summary of the multiple linear regression analysis predicting percentile from left eye dwell time, right eye dwell time, nose dwell time, and mouth dwell time, when outliers are not removed.

Variable	***t***	**b**	**β**	***F***	R^2	**Adj.** R^2	***p***
Overall Model				.781	.078	-.022	.545
Left Eye Dwell Time	-.236	-232.819	-.053				.815
Right Eye Dwell Time	-.669	-421.179	-.155				.508
Nose Dwell Time	.679	227.835	.117				.501